

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 19:45:28 ; Search time 1827 Seconds

(without alignments)

3403.978 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 384

Sequence: 1 MDPAGSGRGVLPKRCRLVL.....CVPSPSKMPQMPPEEPL 384

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308016

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame_p2n.model -DEV-xlh
-O/cgn2_1/USPTO.spool/US09784810/tunat_11072003_103504_27206/app-query.fasta.1.563
-DB-EST -QFWT-fastap -SUFFIX-olig.rst -MINMATCH-0.1 -LOOPEXT-0 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdi -LIST-45
-DOCLALIGN-200 -THR.SCORER-quality -THR.MIN-1 -ALIGN-15 -MODE-LOCAL -OUTFWT-plo
-USER-US09784810.ecgn.1.11906.6tunat_11072003_103504_27206 -NCPU-6 -ICPU-3
-NO_MMAP -LARGESIZE -NEG.SCORER-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV.TIMEOUT-120 -WARN.TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -Fgapop-6
-Fgapext-7 -Ygapop-60 -Ygapext-60 -DELOP-6 -DELEXT-7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	144	37.5	2133	11 BC014439	BC014439 Homo sapi
2	126	32.8	1005	11 BM559257	BM559257 AGENCOURT
3	115	29.9	478	12 BE740866	BE740866 601593092
4	115	29.9	603	10 BE275818	BE275818 601121616
5	115	29.9	852	10 BE274434	BE274434 601120471
6	114	29.7	1232	13 BM556915	BM556915 AGENCOURT
7	111	28.9	662	13 BM386957	BM386957 UT-R-CNI-
8	103	26.8	999	14 BQ675531	BQ675531 AGENCOURT
9	97	25.3	793	12 BG678689	BG678689 602624488
10	96	25.0	1100	14 BM916526	BM916526 AGENCOURT
11	96	25.0	1137	14 BM810136	BM810136 AGENCOURT
12	95	24.7	676	10 AW963415	AW963415 EST375488
13	95	24.7	888	14 BQ647377	BQ647377 AGENCOURT
14	95	24.7	1115	13 BM550039	BM550039 AGENCOURT
15	93	24.2	697	13 B1255900	B1255900 602976518
16	91	23.7	618	10 AW662445	AW662445 h127c05.x
17	91	23.7	1054	14 BM808698	BM808698 AGENCOURT
18	86	22.4	534	12 BE904632	BE904632 601498734
19	85	22.1	720	14 BQ109456	BQ109456 Imageqc-7
20	80	20.8	650	12 BG104264	BG104264 602310848
21	80	20.8	772	9 A1769914	A1769914 wj30d06.x
22	79	20.6	271	14 D31133	D31133 HUM12618 H
23	79	20.6	348	14 BM709951	BM709951 UT-E-CK1-
24	78	20.3	1134	14 BM808666	BM808666 AGENCOURT
25	76	19.8	856	13 B1757316	B1757316 603028531
26	75	19.5	76	9 A0134361	A0134361
27	74	19.5	973	10 BE273426	BE273426 601142888
28	74	19.3	809	12 BG280830	BG280830 602401209
29	73	19.0	785	12 BG680521	BG680521 602628344
30	69	18.0	702	10 BE464487	BE464487 h217910.x
31	65	16.9	318	13 B1013436	B1013436 CM4-ET023
32	61	15.9	356	9 AA081152	AA081152 zn34a06.r
33	61	15.9	386	14 BM797761	BM797761 K-EST0081
34	59	15.4	892	13 B1860351	B1860351 603387479
35	57	14.8	1075	13 BM557357	BM557357 AGENCOURT
36	56	14.6	587	9 AU147968	AU147968 AU147968
37	55	14.3	447	9 A1042283	A1042283 oy13e08.x
38	54	14.1	662	12 BE893504	BE893504 601438057
39	53	13.8	914	13 B1823329	B1823329 603041240
40	52	13.5	255	12 BE930311	BE930311 MR2-NT013
41	47	12.2	671	10 AV694791	AV694791 AV694791
42	44	11.5	206	13 BG987849	BG987849 CM3-HT115
43	43	11.2	341	9 AA026479	AA026479 zj99b03.r
44	43	11.2	604	9 AU122376	AU122376 AU122376
45	43	11.2	875	13 BM006005	BM006005 603613736

ALIGNMENTS

RESULT 1
BC014439 2133 bp mRNA linear HTC 19-DEC-2001
LOCUS BC014439
DEFINITION Homo sapiens, clone IMAGE:4871343, mRNA.
ACCESSION BC014439
VERSION BC014439.1 GI:17939597
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2133)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-2001) National Institutes of Health, Mammalian

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue procurement: ATCC/DCMP/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Jacqueline
Ness, Pavan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schell, Duane Smalish, Michael Smith, Lorralne Spence, Jeff Stolt,
Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 34 Row: 1 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA 91: 11464966
This clone has the following problem: incomplete processing.

FEATURES
source

1. 2133

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4871343"

/tissue_type="Skin, melanotic melanoma, high MDR."

BASE COUNT 350 a 681 c 383 t
ORIGIN

Alignment Scores:

Pred. No.: 5,58e-131 Length: 2133
Score: 144.00 Matches: 375
Percent Similarity: 95.42% Conservative: 0
Best Local Similarity: 95.42% Mismatches: 9
Query Match: 37.50% Indels: 18
DB: 11 Gaps: 0

US-09-784-810A-2 (1-384) x BC014439 (1-2133)

OY 1 MetAspProAlaGlyIysGlyProArgIValLeuProArgProCysArgValLeuValLeu 20
Db 693 ATGGATCCAGCGGGGCGGCCCGGGGGGTGCTCCCGGGCGCTGCGGTGCTG 752
OY 21 LeuAsnProArgGlyIysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
Db 753 CTGAACCCCGCGGGGCGCAAGGCAAGCCCTTGACGCTCTCCGAGTCAAGCGCAAGCC 812
OY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
Db 813 CTTTGGCTGAGGCTGAATCTCTCTTCAAGTGTCACTGAGCGGCGAACCACGCG 872
OY 61 ArgGluLeuValArgSerGluIleuGlyArgGtrPaspAlaLeuValMetSerGly 80
Db 873 CGGAGACTGGTGGCGGAGAGAGCTGGGCGCTGAGCGCTGAGTGTCTGGA 932
OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
Db 933 GACGGCTGATGACAGGAGTGTGGAACGGGCTCAAGGAGCGGCTGAGTGGAGACCC 992
OY 101 IleGlnLysProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
Db 993 ATCAGAAAGCCCGTGTAGCTTCCAGCAGAGGCTTGAGCAAGCGCTGAGCTTCTTG 1052

OY 121 AsnHisTyrAlaGlyIysGluGlnValThrAsnGluAspLeuThrAsnCysThrLeu 140
Db 1053 AACCATTTGCTGCTATGAGCAGGTACCAATGAAGACCTCTGACCAATGCACGCTTA 1112
OY 141 LeuLeuCysArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerG 160
Db 1113 TTGCTGTCCG-CCGGC-TGCTGTACACCATGAACCTGCTGTGACACAGGCTTCGG 1170
OY 160 LysLeuArgSer-PheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValLeu 179
Db 1171 GGCTGGGCT-CTTCTGTCTCTACAGCTGGGCTGGGCTTCAATGCTGTGGACCTTA 1229
OY 180 GluSerAsp-LysTyrArgArgLeuGlyLysMetArgPheThrLeuGlyThrPheLeu 199
Db 1230 GAGAGTGA-GAAGTATCGGCGCTGCGGAGATGGCTTCACTTGGGACCTTCTGCG 1288
OY 199 GluAlaAlaLeuAlaGlyThrTyrArgIValLeuAlaThr-LeuProValGlyArgValG 219
Db 1289 TCTGCGAGCCCTGCGACCTACCGGCGGCGGAGCTGGCTTA-CTTCCCTGTAGGAAGCT-G 1346
OY 219 LysPhe-LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeu 238
Db 1347 GGTTCCAAGACACCTGCTCCCGCTTGTGTCCAGCAGGCGCGGTAGATGCACACTT 1406
OY 239 ValProLeuGluGluGln-ValProSerHisTrpGln-ValValProAspGluAspPhe 258
Db 1407 GTGCCACTGGAGAGGCC-AGTGGCTCTCACTGAGAC-AGTGGTCCCGCAGGAGACTTGG 1464
OY 258 AluLeuValLeuAlaLeuLeuHisSerHisLeuAla-SerGluMetPheAlaIleProMet 277
Db 1465 TGTAGTCTGTGCACTGCTGCACCTCCGACCTCGG-CAAGTGAATGTTGCTCACCCATG 1523
OY 278 GlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAla 297
Db 1524 GGCCCTGTGAGCTGGGCGGTATCATCTGTTCAGTGGCGGCGGAGATGCTGCTGGC 1583
OY 298 MetLeuLeuArgLeuPheLeuAlaMetGlyLysGlyArgHisMetGluIlyrGluCysPro 317
Db 1584 ATGCTGTGCGGCTCTTCTCGGCATGAGAGGAGGCAATGAGATGAAATGCC 1643
OY 318 TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyIysGlyVal 337
Db 1644 TACTTGTGTATGAGTCCCGCTGCTGCTGCTGCTGAGCCCAAGATGGGAAAGGTGG 1703
OY 338 PheAlaValAspGlyIysLeuLeuMetValSerGluAlaValGlnGlyValHisProAsn 357
Db 1704 TTTCGAGTGGATGGGAATGATGATTAGCTGAGCGGCGCTGCAAGGCGAGTGCACCAAC 1763
OY 358 TyrPheTrpMetValSerGlyCysValGluProProSerTrpLysProGlnGlnMet 377
Db 1764 TACTTGTGATGTGTACAGGCTTGTGCGAGCCCGCCAGCTGAGAGCCCAAGATG 1823
OY 378 ProProGluGluProLeu 384
Db 1824 CCACCGGCACAGAGCCCTTA 1844

RESULT 2
BM559257 1005 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6598353 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:547413
DEFINITION 5', mRNA sequence.
ACCESSION BM559257
VERSION BM559257.1 GI:18802694
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1005)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM1987 row: j column: 14
High quality sequence stop: 618.

FEATURES

source

Location/Qualifiers

1. 1005
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5474413"
/clone_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(C). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 157 a 333 c 348 g 167 t
ORIGIN

Alignment Scores:

Pred. No.: 1.43e-113 Length: 1005
Score: 126.00 Matches: 126
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.81% Indels: 0
DB: 13 Gaps: 0

US-09-784-810A-2 (1-384) x BM559257 (1-1005)

OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 326 ATGGATCCAGCGGGCGGGCCCGGCGTCTCCCGCGCCCTCGCTCGTCTG 385
OY 21 LeuAsnProArgGlyGlyGlyValAlaLeuGlnLeuPheArgSerHisValGlnPro 40
DB 386 CTGAACCCCGCGCGCGGCAAGGCGCTTGCAGCTTCCGAGACAGCTGACGCC 445
OY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
DB 446 CTTTGGCTGAGCTGAATCTCTTCACGCTGATGCTCACTGAGCGCGGACACACGCG 505
OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValAlaMetSerGly 80
DB 506 CGGAGCTGGTGGCGGCGAGAGAGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCA 565
OY 81 AspGlyLeuMetHisGluValAlaAsnGlyLeuMetGluArgProAspTrpGluThraAla 100
DB 566 GACGGCGCTGATGCACGAGAGTGTGAACGGCTCATGAGCGCGCTGACTGGAGACGCC 625
OY 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
DB 626 ATCCAGAAAGCCCTGTGTAGCTCCACAGCAGCTCTGCAACGCGCTGCGCTTCCCTTG 685
OY 121 AsnHisTyrAlaGlyTyr 126
DB 686 AACCATTAATGCTGGCTTA 703

RESULT 3 478 bp mRNA linear EST 15-SEP-2000
LOCUS BE740866
DEFINITION 601593092P1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3947182 5',
mRNA sequence.
ACCESSION BE740866

VERSION BE740866.1 GI:10154858
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 478)

AUTHORS

NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov

Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LNCM808 row: g column: 23
High quality sequence stop: 476.

FEATURES

source

Location/Qualifiers

1. 478
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3947182"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(C). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 82 a 134 c 159 g 103 t
ORIGIN

Alignment Scores:

Pred. No.: 4.62e-103 Length: 478
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.95% Indels: 0
DB: 12 Gaps: 0

US-09-784-810A-2 (1-384) x BE740866 (1-478)

OY 270 SerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGlyValMetHisLeuPheTyr 289
DB 125 AGTGAATGTTTCTGTCACCCATGAGCGCGTGTGACGCTGCATGATCTGTTCTAC 184
OY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGly 309
DB 185 GTGCGGGCGGAGTGTCTCGTGCATGCTGCTGCGCGCTTCTCGGCCATGAGAGAGGC 244
OY 310 ArgHisMetGluTyrGluCysProTyrLeuValTyrValProValValAlaPheArgLeu 329
DB 245 AGGCATATGAGATGATGATGCCCTACTTGTGTAATGAGCCCGGTGCTCCGCTTG 304
OY 330 GluProLysAspGlyLysGlyValAlaPheAlaValAspGlyGluLeuMetValSerGluVal 349
DB 305 GAGCCCAAGATGAGAAAGTGTGTTCAGTGTGAGTGGGAAATGATGATGAGGAGGCC 364
OY 350 ValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro 369
DB 365 GTGCAGGGCCAGGCGACCAACTACTTGTGATGTCACGCGGTGCGTGGAGCCCGC 424
OY 370 ProSerTrpLysProGlnGlnMetProProProGluGluProLeu 384
DB 425 CCCAGCTGGAAGCCCAAGATGCCACAGCCGCAAGAGCCCTTA 469

```

RESULT 4
BE275818 603 bp mRNA linear EST 13-JUL-2000
LOCUS 601121616F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3346011 5',
DEFINITION mRNA sequence.
ACCESSION BE275818
VERSION BE275818.1 GI:9150778
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 603)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LNCM134 row: c column: 04
High quality sequence stop: 599.
FEATURES
Source
1. 603
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3346011"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 100 a 167 c 206 g 130 t

ORIGIN

Alignment Scores:

Pred. No.:	5,92e-103	Length:	603
Score:	115.00	Matches:	115
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	29,95%	Indels:	0
DB:	10	Gaps:	0

US-09-784-810a-2 (1-384) x BE275818 (1-603)

QY 270 SeqGluMethPheAlaProMetGlyAArgCysAlaAlaGlyValMetHisLeuPheTyr 289

DB 102 AGTGAGATGTTGGTCGACACCATGGCCGCTGTCACACTGGCTGCGCTATGATGTTCTTAC 161

QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGly 309

DB 162 GTGGGGGGGGAGTGTCTGTCACAGCTGCTGGCTCTCTCTGCGCAGTGGAGAGGCG 221

QY 310 ARGHisMetGluTyrGluCysProTyrLeuValTyrValProValAlaAlaPheArgLeu 329

DB 222 AGGCATATGAGATGATCAATGACCCCTACTGTTATACGTCGCCCTGCTCCCTCCGCTTG 281

QY 330 GlnProLysAspGlyLysGlyValPheAlaValAlaAspGlyLysLeuMetValSerGlnAla 349

DB 282 GAGCCCAAGAGATGGAAAGCTGTCTTTCACAGTGGATGGGAATTTAGTTAGCGAGGCC 341

QY 350 ValGlnGlyGlnValHisProAsnTyrPheTyrMetValSerGlyLysValGlnProPro 369

```

DB 342 GTCCAGGAGCCAGGTGCACCCAACTACTCTGTGATGTCAGCGGTGCGTAGACCCCG 401
QY 370 ProSerTyrLysProGlnGlnMetLeuProProGluGluProLeu 384
DB 402 CCCAGCTGGAGAGCCCGACAGATGTCACACCGCGAAGAGACCCCTTA 446
RESULT 5
BE274434 852 bp mRNA linear EST 13-JUL-2000
LOCUS 601120471F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:2967353 5',
DEFINITION mRNA sequence.
ACCESSION BE274434
VERSION BE274434.1 GI:9149375
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 852)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LNCM17 row: m column: 18
High quality sequence stop: 827.
FEATURES
Source
1. 852
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2967353"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 132 a 278 c 301 g 141 t

ORIGIN

Alignment Scores:

Pred. No.:	8.56e-103	Length:	852
Score:	115.00	Matches:	128
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	29,95%	Indels:	0
DB:	10	Gaps:	0

US-09-784-810a-2 (1-384) x BE274434 (1-852)

QY 1 MetAspProAlaGlyLysLysProArgGlyValLeuProArgProCysArgValLeuValLeu 20

DB 310 ATGATTCACAGGGGGGCGCCCGGGGGCTGCTCCCGGGGCTGCGGCTGCTGGAGCTG 369

QY 21 LeuAsnProArgGlyLysLysLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40

DB 370 CTGAACCCCGCGCGCGCAAGGAGCCCTTTCAGCTCTCCGAGTACCTCCAGGCC 429

QY 41 LeuLeuAla-GlnAlaGlnIleSerPheThrLeuMetLeuThrGlnArgArgAsnHisAl 60

DB 430 CTTTGGCTTGAGCGTGAATATCTCTTACACCTATGCTACTGAGACGGCGGAACACAGC 489


```

Oy 60 aatggtluevalarserglueuglyargtrpaspalaleuvalmetsergl 80
    |||
Db 490 GCGGAGCTGGTGTGCGAGAGAGCGGCGCTGGAGCGCTGTGGTGTCTGCG 549
Oy 80 yaagcglyleumethisgluvalaasnnglyleumecgluarprpsprtgltutnal 100
    |||
Db 550 AGAGGGGCTGATGACAGAGGTGTGAACGGGCTCATGAGCGGCTGTGAGAGCGCG 609
Oy 100 aialgllysrproleucysserleuprolaaglyserglyasnalaaleualasertle 120
    |||
Db 610 CATCAGAAAGCCCTGTGTACCTCCAGAGGCTGTGGCAAGCGCTGCGAGCTTCTT 669
Oy 120 uasnhtyralaaglytyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 128
    |||
Db 670 GAACCATTTATGCTGTGCTATGAGCAG 694

RESULT 6
BM556915 1232 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT.6577988 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467608
DEFINITION 5', mRNA sequence.
ACCESSION BM556915
VERSION BM556915.1 GI:18798522
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1232)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: rgs@bbs.fda.nih.gov
            Tissue Procurement: DCTD/DP
            CDNA Library Preparation: Rubin Laboratory
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            http://image.llnl.gov
            Plate: LUCM1969 row: 0 column: 01
            High quality sequence stop: 447.
FEATURES
    source
        1..1232
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_image="5467608"
            /clone_lib="NIH_MGC_41"
            /tissue_type="amelanotic melanoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCAGAGAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH_MGC Library."
BASE COUNT 250 a 369 c 392 g 220 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.23e-101 Length: 1232
Score: 114.00 Matches: 114
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.69% Indels: 0
DB: 13 Gaps: 0
US-09-784-810A-2 (1-384) x BM556915 (1-1232)
Oy 271 GtmetphealaalalPrometclYARGcYsAlaAlaaglyValMetHisLeuphenYVal 290

```

```

Db 27 GAGATGTTTGTGTCACCCATGGCGGCTGTCACATGCGCATGATCTTCTACGTG 86
Oy 291 ArgAlaGlyValaSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluGlyArg 310
    |||
Db 87 CGGGGGAGAGTGTCTGCGCATCTGTGTGCGCTCTTCTCGGCATGAGAGAGGAGCG 146
Oy 311 HisMetGluTyrgLucysProTyrlleuValTyrlaProValAlaAlaPheArgLeuGlu 330
    |||
Db 147 CATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 206
Oy 331 ProLysaspGlyLysGlyValaPheAlaValaSpLygluLeuMetValSerGluAlaVal 350
    |||
Db 207 CCCAAGGATGGGAAGGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 266
Oy 351 GlnGlyGlnValHisProAsnTyrlPheTrpMetValSerGlyCysValGluProProPro 370
    |||
Db 267 CAGGCGCAGGTGACCCCAAACTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 326
Oy 371 SerTrpLysProGlnGlnMetProProProGluGluProLau 384
    |||
Db 327 AGCTGGAAAGCCCGACAGATGCCACGCCAGAAAGCCCTTA 368

RESULT 7
BM386957/c 662 bp mRNA linear EST 17-JAN-2002
LOCUS UI-R-CNI-cj1-c-13-0-UI-s1 UI-R-CNI Rattus norvegicus cDNA clone
DEFINITION UI-R-CNI-cj1-c-13-0-UI 3', mRNA sequence.
ACCESSION BM386957
VERSION BM386957.1 GI:18187010
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE 1 (bases 1 to 662)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL discovery
MEDLINE Genome Res. 6 (9), 791-806 (1996)
COMMENT 9704447
        Program for Rat Gene Discovery and Mapping
        University of Iowa
        451 Eckstein Medical Research Building Iowa City, IA 52242, USA
        Contact: Soares, MB
        Tel: 319 335 8250
        Fax: 319 335 9565
        Email: mscoares@iue.uiowa.edu
        The sequence contained an oligo-dT track that was present in the
        oligonucleotide that was used to prime the synthesis of first A
        strand cDNA and therefore this may represent a bonafide poly A
        tail. CDNA Library Preparation: M.B. Soares Lab Clone distribution:
        Seq primer: M13 Forward
        POLYA=Yes.
FEATURES
    source
        1..662
            Location/Qualifiers
            /organism="Rattus norvegicus"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /clone_image="UI-R-CNI-cj1-c-13-0-UI"
            /clone_lib="UI-R-CNI"
            /dev_stage="adult"
            /lab_host="DH10B (Life Technologies)"
            /note="Vector: pRT3D-Pac (Pharmacia) with a modified
            polylinker. Site_1: Not I; Site_2: Eco RI; The UI-R-CNI
            library is a subtracted library derived from the following
            pool of seven normalized rat libraries: normalized rat
            seminal vesicles, normalized rat penis, normalized rat
            bladder, normalized rat cervix, normalized rat brown
            adipose, normalized rat fundus, and normalized rat
            salivary gland. It was constructed according to the

```

procedure described by Bonaldi, Lemon & Coates (Genome Research Genome 6: 791-806, 1996). For construction of the CN1 library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CA1 corresponding to plates R-CA0-AWV through R-CA0-AXS, R-CA0-AXX through R-CA0-BAL, R-CA0-BFE through R-CA0-BH, R-CA0-BUS, R-CA0-BKS, R-CA0-BKG-H, R-CA0-BKT-K, R-CA0-BKP through R-CA0-BKC, R-CA0-BKV-U, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BPA through R-CA0-BPG, R-CA1-BBA through R-CA1-BD, R-CA1-BHZ through R-CA1-BIP, R-CA1-BIR, R-CA1-BJ, R-CA1-BKZ through R-CA1-BKB, R-CA1-BKD, R-CA1-BKE, R-CA1-BKI, R-CA1-BKT, R-CA1-BLF, R-CA1-BLH through R-CA1-BLN, R-CA1-BLS, R-CA1-BLV-U, R-CA1-BNR, and the R-CA1-BLE. The resulting final driver population represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CUDs, CWDs, CX0s and normalized libraries CS0, CT0, CUD, CWO, and CX0 corresponding to plates R-CT0s-CAY, R-CUDs-CBP through R-CT0s-CAM through R-CT0s-CAY, R-CUDs-CBM through R-CUDs-CCA, R-CUDs-CCB through R-CUDs-CCM, R-CUDs-CCN through R-CUDs-CCX, R-CS0-BSD, R-CS0-BPD through R-CS0-BTV through R-CX0s-CCX, R-CS0-BSD, R-CS0-BVP, R-CS0-BVN, R-CS0-BWA, R-CT0-BTW through R-CT0-BUP, R-CS0-BVN, R-CUD-BUQ through R-CUD-BVL, R-CMO-BVV through R-CMO-BMP, R-CMO-BXN through R-CMO-BXO, R-CX0-BMQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CVO-BBH through R-CVO-BRR, R-CVI-BRS through R-CVI-BSC, R-CVI-BSE through R-CVI-BTC, and R-CVI-BVO through R-CVI-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, represented about 5% of the final driver population. f) A pool of about 10,000 clones from rat BVO and BVOP (7-9.5 kb DNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BPB, R-BVO-ANK through R-BVO-ANR, R-BVOP-AOI through R-BVOP-AOX, and R-BX0-AQY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer including CY0, CZ0, DA0, DC0, DDO, and DE0 including R-CYO-BXP through R-CYO-BXZ, R-CZO-BYA through R-CZO-BVI, R-CZO-BZC-X, R-DA0-BVJ through R-DAO-BXA, R-DAO-BZD through R-DAO-BZH, R-DB0-BYG through R-DCO-BEA, R-DCO-BZI through R-DCO-BEO, R-DCO-CAY through R-DCO-CGA, R-DD0-BZR through R-DDO-CAA, through R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAI. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CNO, that makes up the tracer. The corresponding plates are R-CNO-BKW through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLM-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CNO pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bkz-c-06-0-UI, bkz-d-09-0-UI, bkz-e-06-0-UI, bkx-a-11-0-UI, bkz-c-09-0-UI, bkz-d-10-0-UI, bla-a-01-0-UI, bla-x-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, bli-a-12-0-UI, bli-f-02-0-UI, bli-c-11-0-UI, bli-e-95-0-UI, bli-d-1-08-0-UI, bli-f-02-0-UI, bli-q-04-0-UI, bli-r-a-05-0-UI, bli-f-08-0-UI. This pool represented 5%

of the final driver population. 1) One abundant CMO clone (corresponding to the address bzr-a-11-0-01) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG-Seq-None found"

BASE COUNT	136 a	210 c	178 g	138 t
ORIGIN				
Alignment Scores:				
Pred. NO.:	5.78e-99	Length: 662		
Score:	111.00	Matches: 111		
Percent Similarity:	100.00%	Conservative: 0		
Best Local Similarity:	100.00%	Mismatches: 0		
Query Match:	28.91%	Indels: 0		
DB:	13	Gaps: 0		

HS-09-784-810A-2 (1-384) x BM386957 (1-662)

OY		274	AlalabPromeGtIvArygCSAlaaIagIvaImethIsLeupheTyVaIArglaIagly	293
Dp		622	GCTGACCCATGGGCCCTGTGCAGGCTGGCGTCATCATCTGTTCACCTGGGGGGGA	563
OY		294	VaISerArGaImetLeuLeuAglLeuPhyLeuAlMeGlLuPySgIvArghISmetGIu	313
Dp		562	GTGTCTCTGCCAATGCTGCTGGCCTTCCCTGGCCATTGAGAAAGGAGGCATATGAG	503
OY		314	TyrGIuCySProTYrLeuValTYrVaIProValVaIaPheArGLeuGIuProLySAAP	333
Dp		502	TATGATGCCCCCTACTTGTTATGTGGCCCTGGCTCCCTTCGCTTGAGGCCCAAGAT	443
OY		334	GIuLYSGIvAlPheAlaValASpLySIuLueMetVAlSerGIuaIaValGInGIyGln	353
Dp		442	GGAAAGAGTGTGTTCAGATGGATGGGAATGATGGTTACCGAGGCCGTCAGAGGCCAG	383
OY		354	ValHISproASTYrPheTrpMetVAlSerGIyCSVAIGluProProPtoSerTrPLys	373
Dp		382	GTGCACCACAACtACTTCTTGATGTGTACAGCGGTTCTGTCGAGCCCGCCAGCTGGAAG	323
OY		374	ProGIuImeTrProProGIuGIuProLeu	384
Dp		322	CCCCAGAGATGCCACGCCAGAGAAGCCCTTA	290
RESULT 8				
BQ675531			999 bp	mRNA linear EST 15-JUL-2002
LOCUS				
DEFINITION			AGNCOURT_8196565 NIH_MGC_102 Homo sapiens CDNA clone IMAGE:62561955	
ACCESSION			5', mRNA sequence.	
VERSION			BQ675531	
KEYWORDS			BQ675531.1 GI:21786365	
SOURCE			EST.	
ORGANISM			human.	
			Homo sapiens	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE			1 (bases 1 to 999)	
AUTHORS			NIH-MGC http://mgc.nci.nih.gov/.	
TITLE			National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL			Unpublished (1999)	
COMMENT			Contact: Robert Strausberg, Ph.D. Email: cgaqbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC Clone Distribution Information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: L1CM2408 row: h column: 20 High quality sequence stop: 538.	
FEATURES			Location/Qualifiers	
source			1..999	

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6256195"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

BASE COUNT 175 a 337 c 315 g 171 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 7.01e-91 Length: 999
Score: 103.00 Matches: 103
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.82% Indels: 0
DB: 14 Gaps: 0

US-09-784-810A-2 (1-384) x BG675531 (1-999)

OY 42 LeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAlaArg 61
DB 3 TTGGCTGAGGCTGAAMTCTCTTACGCTGATGCTCAGAGCGCGGACACCGCGG 62
OY 62 GluLeuValArgSerGluLeuGluArgTrpAspAlaLeuValMetSerGlyasp 81
DB 63 GAGCTGGCGCGTGGAGAGAGCTGGCGCTGGAGCTGTGGTATGCTGGAGAC 122
OY 82 GlyLeuMetHisGluValAlaAsnGlyLeuMetGluArgProAspTrpGluThrAlaIle 101
DB 123 GGGCTGATGACAGAGAGTGGTGAACGGGCTCATGAGCGGCTGACTGGAGACGCCATC 182
OY 102 GluArgProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsn 121
DB 183 CAGAGCGCGCTGTAGGCTCCAGCAGGCTGTGCAACGGCTGGAGCTTCTTGAAC 242
OY 122 HisTrpAlaGlyTrpGluGluValThrAsnGluAspLeuThrAsnCysThrLeuLeu 141
DB 243 CATATGCTGGCTATGACAGAGTCAACATGAAAGCTCTGACCAACTGCACGCTATTG 302
OY 142 LeuGlyArg 144
DB 303 CTGTGCCGC 311

RESULT 9
BG678689 793 bp mRNA linear EST 01-MAY-2001
DEFINITION 602624488P1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749573 5',
mRNA sequence.
ACCESSION BG678689
VERSION BG678689.1 GI:13910086
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNMI0602 row: p column: 22
High quality sequence stop: 669.
Location/Qualifiers
1. 793

FEATURES
SOURCE

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749573"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP library."

BASE COUNT 129 a 244 c 255 g 165 t
ORIGIN

Alignment Scores:
Pred. No.: 4.55e-85 Length: 793
Score: 97.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.26% Indels: 0
DB: 12 Gaps: 0

US-09-784-810A-2 (1-384) x BG678689 (1-793)

OY 48 SerPheThrLeuMetLeuThrGluArgAsnHisAlaArgGluLeuValArgSerGlu 67
DB 2 TCTTCACGCTGATGCTACTGAGCGGGAACACCGCGGAGCTGCGGTGGAG 61
OY 68 GluLeuGluArgTrpAspAlaLeuValMetSerGlyAspGlyLeuMetHisGluVal 87
DB 62 GAGCTGGCGCGCTGGAGAGCTGTGGTGTGCTCATGCTGAGACGGGCTGATGACAGAGGTG 121
OY 88 ValAsnGlyLeuMetGluArgProAspTrpGluThrAlaIleGluArgProLeuGlySer 107
DB 122 GTGAACGGGCTCATGAGAGCGGCTGAGAGAGCGGCTGACAGAGCGGCTGTGAGC 181
OY 108 LeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisTrpAlaGlyTrpGlu 127
DB 182 CTCGACAGAGGCTGTGGCAACGCGTGGACGCTTCTTGAACATATGCGGCTATGAG 241
OY 128 GluValThrAsnGluAspLeuThrAsnCysThrLeuLeuGlyArg 144
DB 242 CAGGTACCAATGAAAGACTCTCTGACCAACTGACAGCTATGCTGTGCCGC 292

RESULT 10
BM916526 1100 bp mRNA linear EST 12-MAR-2002
DEFINITION 6641989 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482958
5', mRNA sequence.
ACCESSION BM916526
VERSION BM916526.1 GI:19366905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1100)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue procurement: DCPD/DMP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2009 row: n column: 15
High quality sequence stop: 625.

FEATURES
source

1. 1100
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 187 a 333 c 348 g 222 t 10 others

ORIGIN

Alignment Scores:

Pred. No.: 6,26e-84 Length: 1100
Score: 96.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.00% Indels: 0
DB: 14 Gaps: 0

US-09-784-810A-2 (1-384) x BM916526 (1-1100)

QY 49 PheHleuMetLeuThrgLuarGrAgsnHsAlaArgLgIuLeuValArgSerGluGlu 68
|||||
DB 3 TTCACGCTGATGCTCCTAGCGGGGGAACCAACCGCGGAGCTGCTGGCGAGAG 62
QY 69 LeuGlyArgTrpAspAlaLeuValValMetSerGlyAspGlyLeuMetHisGluVal 88
|||||
DB 63 CTGGGGCGCTGGGAGCCCTGCTGCTCATGCTGAGACGCGCTGATGACAGAGTG 122
QY 89 AsnGlyLeuMetGluArgProAspTrpGluThrAlaIleGlnIysProLeuGlySerLeu 108
|||||
DB 123 AACGGGCTATGGAGGGGCTGACTGGGAGACCGCCATCCAGAACGCCCTGTAGCTC 182
QY 109 ProAlaGlySerGlyAsnAlaLeuAlaIleSerLeuAsnHisTyrAlaGlyTyrGluGln 128
|||||
DB 183 CCAGCAGGCTCTGGCAGCGCTGGCAGCTTCTTGAAACCATTAAGCTGCTATGAGCAG 242
QY 129 ValThrAsnGluAspLeuLeuThrAsnCysThrLeuLeuGlyCysArg 144
|||||
DB 243 GTCACCAATGAGAGCCTCTGACCAATGACGCTATTGCTGTGGCCG 290

RESULT 11

BM810136

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM810136 1137 bp mRNA linear EST 05-MAR-2002
AGENCOURT_6579661 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5470089
5' mRNA sequence.
BM810136
BM810136.1 GI:19126959
EST

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 1137)

NIH-MGC http://imgc.nhl.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: DCID/DP

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1976 row: f column: 10
High quality sequence stop: 623.

FEATURES
source

1. 1137
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 188 a 342 c 363 g 238 t 6 others

ORIGIN

Alignment Scores:

Pred. No.: 6,48e-84 Length: 1137
Score: 96.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.00% Indels: 0
DB: 14 Gaps: 0

US-09-784-810A-2 (1-384) x BM810136 (1-1137)

QY 49 PheHleuMetLeuThrgLuarGrAgsnHsAlaArgLgIuLeuValArgSerGluGlu 68
|||||
DB 3 TTCACGCTGATGCTCCTAGCGGGGGAACCAACCGCGGAGCTGCTGGCGAGAG 62
QY 69 LeuGlyArgTrpAspAlaLeuValValMetSerGlyAspGlyLeuMetHisGluVal 88
|||||
DB 63 CTGGGGCGCTGGGAGCCCTGCTGCTCATGCTGAGACGCGCTGATGACAGAGTG 122
QY 89 AsnGlyLeuMetGluArgProAspTrpGluThrAlaIleGlnIysProLeuGlySerLeu 108
|||||
DB 123 AACGGGCTATGGAGGGGCTGACTGGGAGACCGCCATCCAGAACGCCCTGTAGCTC 182
QY 109 ProAlaGlySerGlyAsnAlaLeuAlaIleSerLeuAsnHisTyrAlaGlyTyrGluGln 128
|||||
DB 183 CCAGCAGGCTCTGGCAGCGCTGGCAGCTTCTTGAAACCATTAAGCTGCTATGAGCAG 242
QY 129 ValThrAsnGluAspLeuLeuThrAsnCysThrLeuLeuGlyCysArg 144
|||||
DB 243 GTCACCAATGAGAGCCTCTGACCAATGACGCTATTGCTGTGGCCG 290

RESULT 12

AW63415/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AW63415 676 bp mRNA linear EST 01-JUN-2000
ES3375488 MAGE resequences, MACH Homo sapiens cDNA, mRNA sequence.
AW63415
AW63415.1 GI:8153251
EST

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 676)

Hegde, P., Qi, R., Aherathy, K., Dharap, S., Gaspard, R., Gay, C., Holt

I.E., Seed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and

Quackenbush, J.

Assessment of gene expression patterns in a model of colon tumor

metastasis using a 19,200 element cDNA microarray

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLML2769 row: 1 column: 18
High quality sequence start: 20
High quality sequence stop: 451.
Location/Qualifiers

FEATURES
Source

1.1115
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5745953"
/clone.lib="NIH_MGC_118"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6; Site.1: NotI; Site.2: EcoRV
(destroyed); RNA source: leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."
BASE COUNT 177 a 361 g 385 c 192 t
ORIGIN

Alignment Scores:

Pred. No.: 6.16e-83 Length: 1115
Score: 95.00 Matches: 95
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.74% Indels: 0
DB: 13 Gaps: 0

US-09-784-810a-2 (1-384) x BM550039 (1-1115)

QY 1 MetaspProAlagIyGlyProArGlyValleuProArGProCysArGValleuValleu 20
DB 360 ATGATTCACAGCGGGCGCCCGCGGCTGCTCCCGCGCGCTGCTGCTGCTGCTG 419
QY 21 LeuaspProArGlyGlyLysGlyLysAlaLeuGlnLeuPheArGSerHisValGlnPro 40
DB 420 CTGACCGCGCGCGGCTGAAGGCGCAAGGCTTCAGCTCTTCGCGAGTCACGTGACGCC 479
QY 41 LeuenuAlagIuAluIleSerPheRheLeuMetLeuRguLuarGArgAsnHisAla 60
DB 480 CTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCACTGAGCGCGGCAACACACGCG 539
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
DB 540 CGGAGCTGGGCGCGGAGGAGGAGGAGGCGCGCGGAGGAGGAGGAGGAGGAGGAGG 599
QY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgPro 95
DB 600 GACGGGCTGATGCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 644

RESULT 15

B1255900

LOCUS 602976518F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5115605 5',
DEFINITION mRNA sequence.

ACCESSION

B1255900

VERSION

B1255900.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 697)

AUTHORS

NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLML281 row: 4 column: 06
High quality sequence stop: 688.
Location/Qualifiers

FEATURES

Source

1.697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5115605"
/clone.lib="NIH_MGC_12"
/tissue_type="Cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-Sport6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies. 263 g 94 t
BASE COUNT 110 a 230 c 263 g 94 t
ORIGIN

Alignment Scores:

Pred. No.: 3.51e-81 Length: 697
Score: 93.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.22% Indels: 0
DB: 13 Gaps: 0

US-09-784-810a-2 (1-384) x B1255900 (1-697)

QY 7 ProArGlyValleuProArGProCysArGValleuValleuLeuaspProArGlyGly 26
DB 368 CCCCCGGGCGTGTCCCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
QY 27 LysGlyLysAlaLeuGlnLeuPheArGSerHisValGlnProLeuValAlaGlu 46
DB 428 AAGGCGAAGGCTTCGACGCTCTTCGCGAGTACGTCAGCGCCCTTTGGCTGAGGCTGA 487
QY 47 IleSerPheRheLeuMetLeuRguLuarGArgAsnHisAlaArgGluLeuValArgSer 66
DB 488 ATCTCTTACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547
QY 67 GluGluLeuGlyArgTrpAspAlaLeuValMetSerGlyAspGlyLeuMetHisGly 86
DB 548 GAGGAGCTGGGCGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 607
QY 87 ValValAsnGlyLeuMetGluArgProAspTrpGlu 99
DB 608 GTGGTGAACGGGCTCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 646

Search completed: July 12, 2003, 21:23:07

Job time : 1833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 19:55:33 ; Search time 218 Seconds
(without alignments)
2781.350 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 384
Sequence: 1 MDPAGPRGVLPKRCRLV.....CVPKPPKRPQMPPEEPL 384

Scoring table:

OLIGO	Xgapop 60.0	Xgapext 60.0
	Ygapop 60.0	Ygapext 60.0
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 1105431 segs, 789497651 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2208254

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame_p2n.model -DEV-rlh
-Q-/cgn2_1/USPTO.spool/US09784810/rnat_11072003_103505_27238/app_query.fasta-1.583
-DB-published.Applications_NA-OFTM-fastap -SUFFIX-olig.rnpb -MINMATCH-0.1
-LOOPCL-0 -LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo
-TRANS-human40.ccd -LIST-45 -DOCALIGN-200 -THR.SCORE-quality -THR_MIN-1
-ALIGN-15 -MODE-LOCAL -OUTEXT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0
-MATELEN-200000000 -USER-US09784810.ecgn_1.1.57.ernat_11072003_103505_27238
-NCPU-6 -ICPU-3 -NO_MMAP -LARGEDUERY -NEG_SCORES-0 -WAIT -DSPLOCK-100
-LONGLOC -DEV-TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60
-FGAPOP-6 -FGAPEXT-7 -YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database: Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	384	100.0	1600	10 US-09-784-810A-1
2	144	37.5	1155	10 US-09-970-516-1
3	115	29.9	1533	9 US-10-023-282-90
4	65	16.9	199	9 US-09-796-692-2905

Result No.	Score	Query Match Length	ID	Description
5	65	16.9	199	9 US-10-040-862-2905
6	63	16.4	296	9 US-10-015-219-658
7	63	16.4	296	9 US-09-777-564-658
8	39	10.2	394	10 US-09-954-456-1756
9	33	8.6	1149	10 US-09-970-516-5
10	33	8.6	1759	10 US-09-784-810A-3
11	26	6.3	480	10 US-09-783-590-9248
12	9	2.3	30	10 US-09-796-487-17
13	9	2.3	436	10 US-09-960-352-8686
14	9	2.3	1857	10 US-09-970-516-3
15	9	2.3	2380	10 US-09-817-676A-13
16	8	2.1	339	9 US-09-918-995-36854
17	8	2.1	412	10 US-09-960-352-13161
18	8	2.1	440	10 US-09-960-352-12927
19	8	2.1	459	9 US-09-918-995-36954
20	8	2.1	463	9 US-09-918-995-11171
21	8	2.1	600	10 US-09-864-761-12586
22	8	2.1	948	9 US-10-128-714-2287
23	8	2.1	954	9 US-10-156-761-3934
24	8	2.1	1128	10 US-09-956-004-25
25	8	2.1	1185	9 US-10-128-714-1287
26	8	2.1	1281	9 US-10-128-714-7287
27	8	2.1	1314	9 US-10-198-846-10190
28	8	2.1	1425	9 US-10-267-255-81
29	8	2.1	1506	9 US-10-156-761-5829
30	8	2.1	1530	10 US-09-925-301-512
31	8	2.1	1562	9 US-10-128-714-6287
32	8	2.1	1607	10 US-09-925-302-336
33	8	2.1	1789	10 US-09-840-704-1
34	8	2.1	1792	10 US-09-822-830A-420
35	8	2.1	1822	10 US-09-929-918-8
36	8	2.1	2645	9 US-09-764-891-110
37	8	2.1	2658	9 US-09-764-891-6556
38	8	2.1	2837	9 US-09-822-846-58
39	8	2.1	3185	9 US-10-128-714-287
40	8	2.1	3562	9 US-10-128-714-5287
41	8	2.1	4808	9 US-10-142-174-6
42	8	2.1	4901	10 US-09-764-864-1710
43	8	2.1	5258	9 US-09-914-891-7457
44	8	2.1	5763	12 US-10-002-600-79
45	8	2.1	12332	9 US-09-764-891-9634

ALIGNMENTS

RESULT 1
US-09-784-810A-1
Sequence 1, Application US/09784810A
Patent No. US20020082203A1
GENERAL INFORMATION:
APPLICANT: RASTRELLI, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784, 810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1600
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (1)
OTHER INFORMATION: a, t, c, g, other or unknown
US-09-784-810A-1
Alignment Scores:

EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 881
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 880
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 896
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/049, 020
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 876
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 895
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 884
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 894
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 971
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 964
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 882
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 900
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 899
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 893
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 892
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 915
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/049, 019
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 970
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 972
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 916
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/049, 373
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 875
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/049, 374
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 917
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 949
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 974
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 883
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 897
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 898
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 962
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 963
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 877
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/048, 878
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/070, 923
EARLIER	FILING DATE:	1997-12-18
EARLIER	APPLICATION NUMBER:	60/092, 921
EARLIER	FILING DATE:	1998-07-15

EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 90
LENGTH: 1533
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (112)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (123)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1522)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1527)
OTHER INFORMATION: n equals a,t,g, or c
US-10-023-282-90

Alignment Scores:
Pred. No.: 6,92e-106 Length: 1533
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.95% Indels: 0
Gaps: 0
DB: 9

US-09-784-810A-2 (1-384) x US-10-023-282-90 (1-1533)

QY 270 SerGluMetPheAlaIleProMetGluYrGysAlaIleGlyValMetHisLeuPheTyr 289
DB 881 AGTGAATGTTGGCTGCGACCCATGGGCGGCTGTGACGTCGTCATGTCATCTGTCTAC 940
QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGly 309
DB 941 GTGGGGGGGAGATGTCGTGTCGATGTCGTGCGCTCTCTCCGTCGATGAGAGAGGC 1000
QY 310 ArgHisMetGluTyrGluYrGluYrProTyrLeuValTyrValProValAlaIlePheArgLeu 329
DB 1001 AGCATATGAGATGATGAATGCCCTACTGTGATATGTCGCCGTGTCGCTCCGCTTG 1060
QY 330 GluProLysAspGlyLysGlyValPheAlaValAspGlyLysLeuMetValSerGluAla 349
DB 1061 GAGCCCAAGATGGGAAAGTGTGTCAGTGTGATGGGAAATGTGTTAGCGAGGCC 1120
QY 350 ValGlnGlyValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro 369
DB 1121 GTGAGGGGCCAGGTGACCCCAACTACTTGTGATGTCACGCGTTGCGTGAGGCCCGCC 1180
QY 370 ProSerTrpLysProGlnGlnMetProProGluGluProLeu 384
DB 1181 CCCAGCTGAGAGCCCAAGAGATGCCACGCCGAGAGAGCCCTTA 1225

RESULT 4
US-09-796-692-2905
Sequence 2905, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Mannion, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077,001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2905
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-2905

Alignment Scores:
Pred. No.: 1,95e-56 Length: 199
Score: 65.00 Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.93% Indels: 0
Gaps: 0
DB: 9

US-09-784-810A-2 (1-384) x US-09-796-692-2905 (1-199)

QY 306 MetGluLysGlyArgHisMetGluTyrGluYrGluYrProTyrLeuValTyrValProValAla 325
DB 3 ATGGAGAGGCGAGCATATGAGATGATGAATGCCCTACTGTGATATGTCGCCGTGTCG 62
QY 326 AlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaValAspGlyLysLeuMet 345
DB 63 GCCTCCGCTTGAGGCCCAAGATGGGAAAGTGTGTCAGTGTGATGGGAAATGTGATG 122
QY 346 ValSerGluAlaValGlnGlyValHisProAsnTyrPheTrpMetValSerGlyCys 365
DB 123 GTTAGCGAGGCGCTGACGAGGCCAGTGCACCCCAACTACTTGTGATGTCAGTGTGTC 182
QY 366 ValGluProProPro 370
DB 183 GTGAGGCCCGCGGCC 197

RESULT 5
US-10-040-862-2905
Sequence 2905, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Mannion, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06

```

; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2905
; LENGTH: 199
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-2905

Alignment Scores:
Pred. No.: 1,956-56      Length: 199
Score: 65.00             Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.93%      Indels: 0
DB: 9                    Gaps: 0

```

```

US-09-784-810A-2 (1-384) x US-10-040-862-2905 (1-199)
QY 306 MetGlyArgHisMetGlyTyrGluCysProTyrLeuValTyrValProValAla 325
DB 3 ATGGAGAAAGGCGAGCATATGAGATGATGATGATGATGATGATGATGATGATGATG 62
QY 326 AlaPheArgLeuGluProLysAspGlyLysGlyValAlaPheAlaValAspGlyLeuMet 345
DB 63 GCCTTCCGCTTGAGAGCCCAAGATGGAAGGTGTGTGTCAGTGGATGGGAATTTGATG 122
QY 346 ValSerGluAlaValAlaGlnGlyValAlaHisProAsnTyrPheTyrMetValSerGlyCys 365
DB 123 GTTACGAGAGCGCGTCAGAGGCCAGGTGACCCCAAACTACTTGTGATGTCAGTGTTC 182
QY 366 ValGluProProPro 370
DB 183 GTGGAGCCCGCCGCC 197

```

```

RESULT 6
US-10-015-219-658
; Sequence 658, Application US/10015219
; Publication No. US20030008299A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493C1
; CURRENT APPLICATION NUMBER: US/10/015,219
; CURRENT FILING DATE: 2002-03-02
; NUMBER OF SEQ ID NOS: 1739
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 658
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 218, 229, 235, 263, 265, 271, 286, 289
; OTHER INFORMATION: n = A,T,C or G
US-10-015-219-658

```

```

Alignment Scores:
Pred. No.: 2,836-54      Length: 296
Score: 63.00             Matches: 63
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.41%      Indels: 0
DB: 9                    Gaps: 0

```

US-09-784-810A-2 (1-384) x US-10-015-219-658 (1-296)

```

QY 308 LysGlyArgHisMetGlyTyrGluCysProTyrLeuValTyrValProValAlaPhe 327
DB 10 AAGGCGAGCATATGAGATGAAATGCCCTTGTGATATGTCGCCGTGCTTC 69
QY 328 ArgLeuGluProLysAspGlyLysGlyValAlaPheAlaValAspGlyLeuMetValSer 347
DB 70 GCCTTGAGAGCCCAAGATGGAAGGTGTGTGTCAGTGGATGGGAATTTGATGCTTAC 129
QY 348 GluAlaValAlaGlnGlyValAlaHisProAsnTyrPheTyrMetValSerGlyCysValGlu 367
DB 130 GAGGCCGTCAGAGGCCAGGTGACCCCAAACTACTTGTGATGTCAGCGGTTGCTGGAG 189
QY 368 ProProPro 370
DB 190 CCCCAGCC 198

```

```

RESULT 7
US-09-777-564-658
; Sequence 658, Application US/09777564
; Patent No. US2002022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777,564
; CURRENT FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 1730
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 658
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(296)
; OTHER INFORMATION: n = A,T,C or G
US-09-777-564-658

```

```

Alignment Scores:
Pred. No.: 2,836-54      Length: 296
Score: 63.00             Matches: 63
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.41%      Indels: 0
DB: 10                  Gaps: 0

```

US-09-784-810A-2 (1-384) x US-09-777-564-658 (1-296)

```

QY 308 LysGlyArgHisMetGlyTyrGluCysProTyrLeuValTyrValProValAlaPhe 327
DB 10 AAGGCGAGCATATGAGATGAAATGCCCTTGTGATATGTCGCCGTGCTTC 69

```

Page 6

RESULT 9
US-09-970-516-5
: Sequence 5, Application US/09970516

```

77 ValmetSerglyAspglyLeumethHisgluValalasnIlyLeumetGluarproasp 96
|||||

```

```

Db      555  GTCATGTCGGGTGATGCTGATGCATGACGTGATGGCTAATGACAGCCAGAC 614
Qy      97   TTTCTTTTAAATATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
Db      615  TGGAGAGCTGCCATCCAGAAACCCCTGTGTAGCCTCCCT 653

RESULT 11
US-09-783-590-9248
: Sequence 9248, Application US/09783590
: Patent No. US20020110850A1
: GENERAL INFORMATION:
: APPLICANT: Haseltine, William A.
: APPLICANT: L.I. Haseltine
: APPLICANT: Rosen, Craig A.
: APPLICANT: Ruben, Steven M.
: TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
: FILE REFERENCE: PO-16,2C1
: CURRENT FILING DATE: 2000-07-15
: PRIOR APPLICATION NUMBER: US/09/783,590
: PRIOR FILING DATE: 1995-04-12
: PRIOR APPLICATION NUMBER: 08/346,731
: PRIOR FILING DATE: 1994-11-21
: NUMBER OF SEQ ID NOS: 12485
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9248
: LENGTH: 480
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (68)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (193)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (220)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (221)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (236)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (289)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (297)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (302)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (322)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (337)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (350)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (356)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (367)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (382)
: OTHER INFORMATION: n equals a,t,g, or c

```

```

: NAME/KEY: misc feature
: LOCATION: (384)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (386)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (391)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (400)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (412)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (451)
: OTHER INFORMATION: n equals a,t,g, or c
: OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9248

Alignment Scores:
Pred. No.: 6.45e-17 Length: 480
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.77% Indels: 0
Db: 10 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-783-590-9248 (1-480)
Qy      47   ILeSerPheThrLeuMetLeuThrGluArgArgAsnHisAlaArgGluLeuValArgSer 66
Db      12   ATCTCTTACACCTGATGCTCTACTGAGCGGAGAACCGCGGAACTGTGCGATCG 71
Qy      67   GluGluLeuGlyArgTyr 72
Db      72   GAGGAGCTGCGGCGCGCTGG 89

RESULT 12
US-09-796-487-17/c
: Sequence 17, Application US/09796487
: Patent No. US20020042358A1
: GENERAL INFORMATION:
: APPLICANT: Spiegel, Sarah
: TITLE OF INVENTION: Sphingosin kinase, Cloning, Expression and Methods of Use
: FILE REFERENCE: 07320001a (2033957-0001)
: CURRENT APPLICATION NUMBER: US/09/796,487
: CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: US 60/186,532
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: US 09/530,868
: PRIOR FILING DATE: 2000-05-05
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 17
: LENGTH: 30
: TYPE: DNA
: ORGANISM: Unknown
: FEATURE:
: NAME/KEY: Artificial sequence of the antisense primer sphk1-GSP5 in line
: OTHER INFORMATION: , page 8.
: NAME/KEY: misc-feature
: LOCATION: (1)-(30)
: OTHER INFORMATION: Corresponding to the artificial sequence of antisense primer sp
: OTHER INFORMATION: -GSP5 in line 5-6, page 8.
US-09-796-487-17

Alignment Scores:
Pred. No.: 0.67 Length: 30
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.34% Indels: 0

```

DB: 10 Gaps: 0

US-09-784-810a-2 (1-384) x US-09-796-487-17 (1-30)

QY 376 GImetpProProgluGlutProleu 384
 |||||||||||||||||||

Db 30 CAGATGCACCGCCAGAGGCCCTTA 4

RESULT 13

US-09-960-352-8686
 ; Sequence 8686, Application US/09960352
 ; Patent No. US20020137139A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.
 ; APPLICANT: Tao, Ningling
 ; APPLICANT: Byatt, John C.
 ; APPLICANT: Mathialagan, Nagappan
 ; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 ; FILE REFERENCE: 16511.006/37-21(10298)C
 ; CURRENT FILING DATE: 2001-09-24
 ; NUMBER OF SEQ ID NOS: 15112
 ; SEQ ID NO 8686
 ; LENGTH: 436
 ; TYPE: DNA
 ; ORGANISM: Bos taurus
 ; OTHER INFORMATION: Clone ID: 37-LIB34-039-Q1-E1-B2
 ; US-09-960-352-8686

Alignment Scores:

Pred. No.:	7.11	Length:	436
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.34%	Indels:	0
DB:	10	Gaps:	0

US-09-784-810a-2 (1-384) x US-09-960-352-8686 (1-436)

QY 64 ValArgSerGluLeuGlyArgTrp 72
 |||||||||||||||||||

Db 6 GTCCGAGGTAGAGACTCGGCGCTGG 32

RESULT 14

US-09-970-516-3
 ; Sequence 3, Application US/09970516
 ; Patent No. US20020099029A1
 ; GENERAL INFORMATION:
 ; APPLICANT: No. US20020099029A1artis Ag
 ; TITLE OF INVENTION: Induction of blood vessel formation through administration of
 ; FILE REFERENCE: 4-31617
 ; CURRENT FILING DATE: 2001-10-04
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1857
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1857)
 ; OTHER INFORMATION:
 ; US-09-970-516-3

Alignment Scores:

Pred. No.:	25.5	Length:	1857
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.34%	Indels:	0

DB: 10 Gaps: 0

US-09-784-810a-2 (1-384) x US-09-970-516-3 (1-1857)

QY 299 LeuLeuArgLeuPheLeuAlaMetGlu 307
 |||||||||||||||||||

Db 1630 CTGCTGCGCCTTTTCTTGCCATGAG 1656

RESULT 15

US-09-817-676a-13
 ; Sequence 13, Application US/09817676A
 ; Patent No. US20020042101A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Spiegel, Sarah
 ; APPLICANT: Kohama, Takafumi
 ; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
 ; FILE REFERENCE: 00170/HG
 ; CURRENT FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: US 60/194,318
 ; PRIOR FILING DATE: 2000-04-03
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 13
 ; LENGTH: 2380
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (7)..(1860)
 ; PUBLICATION INFORMATION:
 ; TITLE: Molecular cloning and functional characterization of a
 ; JOURNAL: J. Biol. Chem.
 ; VOLUME: 275
 ; ISSUE: 26
 ; PAGES: 19513-19520
 ; DATE: 2000-06-30
 ; DATABASE ACCESSION NUMBER: AF245447
 ; DATABASE ENTRY DATE: 2000-06-27
 ; US-09-817-676a-13

Alignment Scores:

Pred. No.:	31.8	Length:	2380
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.34%	Indels:	0
DB:	10	Gaps:	0

US-09-784-810a-2 (1-384) x US-09-817-676a-13 (1-2380)

QY 299 LeuLeuArgLeuPheLeuAlaMetGlu 307
 |||||||||||||||||||

Db 1636 CTGCTGCGCCTTTTCTTGCCATGAG 1662

Search completed: July 12, 2003, 21:28:09
 Job time : 223 secs

APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 18609 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-1

Alignment Scores:
Pred. No.: 125 Length: 18609
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.34% Indels: 0
Gaps: 0
DB: 4

US-09-784-810A-2 (1-384) x US-08-943-731-1 (1-18609)
QY 104 ProLeucySerLeuProAlaGlySer 112
DB 9033 CCCCTTGCTCTCCTCCAGAGGTC 9059

RESULT 2
US-08-976-259-25
Sequence 25, Application US/08976259
Patent No. 6316609
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Choi, Gil H.
APPLICANT: Welch, Rodney A.
TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
Patent No. 6316609
NUMBER OF SEQUENCES: 142
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,259
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488, 0740002/EKS/CBM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 1128 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-976-259-25

Alignment Scores:
Pred. No.: 92.8 Length: 1128
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
Gaps: 0
DB: 4

US-09-784-810A-2 (1-384) x US-08-976-259-25 (1-1128)
QY 162 ArgSerPheSerValLeuSerLeu 169
DB 857 AGATCTTCAGTCTCCTCGAGTCTT 880

RESULT 3
US-08-770-544-3/C
Sequence 3, Application US/08770544
Patent No. 5907085
GENERAL INFORMATION:
APPLICANT: Gonsalves, Dennis
APPLICANT: Ling, Kai-Shu
TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
THEIR USES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/770,544
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60009008
FILING DATE: 21-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/621
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1602 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-770-544-3

Alignment Scores:
Pred. No.: 128 Length: 1602
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
Gaps: 0
DB: 2

US-09-784-810A-2 (1-384) x US-08-770-544-3 (1-1602)

QY 203 LeuArgThrTyArgGlyArgLeu 210

Db 646 TTACGACCTATCGTGTGACTA 623

RESULT 4

US-09-035-706-1/C

; Sequence 1, Application US/09035706

; Patent No. 6001622

; GENERAL INFORMATION:

; APPLICANT: Dedhar, Shoukat

; APPLICANT: Hannigan, Greg

; TITLE OF INVENTION: Integrin-Linked Kinase and

; TITLE OF INVENTION: Its Uses

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bozicevic & Reed, LLP

; STREET: 285 Hamilton Avenue, Suite 200

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/035,706

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J

; REGISTRATION NUMBER: 36,677

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-327-3400

; TELEFAX: 650-327-3231

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1789 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

US-09-035-706-1

Alignment Scores:

Pred. No.: 142 Length: 1789

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.08% Indels: 0

DB: 3 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-035-706-1 (1-1789)

QY 161 LeuArgSerPheSerValLeuSer 168

Db 239 TTGAGTCGTCTCGTGTGTGCC 216

RESULT 5

US-08-955-841-1/C

; Sequence 1, Application US/08955841

; Patent No. 6013782

; GENERAL INFORMATION:

; APPLICANT: Dedhar, Shoukat

; APPLICANT: Hannigan, Greg

; TITLE OF INVENTION: Integrin-Linked Kinase and

; TITLE OF INVENTION: Its Uses

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bozicevic & Reed, LLP

; STREET: 285 Hamilton Avenue, Suite 200

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/955,841

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: KIN-2C1P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-327-3400

; TELEFAX: 650-327-3231

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1789 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: CDNA

US-08-955-841-1

Alignment Scores:

Pred. No.: 142 Length: 1789

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.08% Indels: 0

DB: 3 Gaps: 0

US-09-784-810A-2 (1-384) x US-08-955-841-1 (1-1789)

QY 161 LeuArgSerPheSerValLeuSer 168

Db 239 TTGAGTCGTCTCGTGTGTGCC 216

RESULT 6

US-09-428-219-3/C

; Sequence 3, Application US/09428219

; Patent No. 6177273

; GENERAL INFORMATION:

; APPLICANT: C. Frank Bennett

; APPLICANT: Lex M. Cowsett

; TITLE OF INVENTION: ANTISENSE MODULATION OF INTEGRIN-LINKED KINASE EXPRESSION

; FILE REFERENCE: RTS-0101

; CURRENT APPLICATION NUMBER: US/09/428,219

; CURRENT FILING DATE: 1999-10-27

; NUMBER OF SEQ ID NOS: 89

; SEQ ID NO: 3

; LENGTH: 1789

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (157)..(1515)

```
US-09-428-219-3
Alignment Scores:
Pred. No.: 142      Length: 1789
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.08%      Indels: 0
DB: 4      Gaps: 0

US-09-784-810A-2 (1-384) x US-09-428-219-3 (1-1789)
QY 161 LeuArgSerPheSerValIleuSer 168
DB 239 TTGAGTCGCTTCCTCCGCTGTGCC 216

RESULT 7
US-09-390-425-1/C
; Sequence 1, Application US/09390425
; Patent No. 6338958
; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; APPLICANT: Hannigan, Greg
; TITLE OF INVENTION: Interglin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/390,425
; EARLIER FILING DATE: 1999-09-03
; EARLIER APPLICATION NUMBER: US60/009,074
; EARLIER FILING DATE: 1995-12-21
; EARLIER APPLICATION NUMBER: US08/752,345
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 1
; LENGTH: 1789
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)...(1512)
; NAME/KEY: other
; LOCATION: (0)...(0)
US-09-390-425-1

Alignment Scores:
Pred. No.: 142      Length: 1789
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.08%      Indels: 0
DB: 4      Gaps: 0

US-09-784-810A-2 (1-384) x US-09-390-425-1 (1-1789)
QY 161 LeuArgSerPheSerValIleuSer 168
DB 239 TTGAGTCGCTTCCTCCGCTGTGCC 216

RESULT 8
US-09-566-906-1/C
; Sequence 1, Application US/09566906
; Patent No. 6369205
; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; APPLICANT: Hannigan, Greg
; TITLE OF INVENTION: Interglin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/566,906
; EARLIER FILING DATE: 2000-05-09
; EARLIER APPLICATION NUMBER: 09/390,425
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: US08/752,345
; PRIOR FILING DATE: 1996-11-19

; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 1
; LENGTH: 1789
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)...(1512)
; NAME/KEY: other
; LOCATION: (0)...(0)
US-09-566-906-1

Alignment Scores:
Pred. No.: 142      Length: 1789
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.08%      Indels: 0
DB: 4      Gaps: 0

US-09-784-810A-2 (1-384) x US-09-566-906-1 (1-1789)
QY 161 LeuArgSerPheSerValIleuSer 168
DB 239 TTGAGTCGCTTCCTCCGCTGTGCC 216

RESULT 9
US-08-321-686B-1
; Sequence 1, Application US/08321686B
; Patent No. 6153427
; GENERAL INFORMATION:
; APPLICANT: King, Donna
; APPLICANT: Schneider, Gary B.
; TITLE OF INVENTION: Treatment of
; TITLE OF INVENTION: Osteopenias and No. 6153427-Human Transgenic
; TITLE OF INVENTION: Mammals Expressing Therapeutic
; TITLE OF INVENTION: Transgenes
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wallenstein & Wagner, Ltd.
; STREET: 311 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch
; MEDIUM TYPE: 800 Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Wordperfect 5.1
; SOFTWARE: Simple Text ASCII (IBM format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,686B
; FILING DATE: October 12, 1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA: not
; PRIOR APPLICATION DATA: applicable
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry, Alan L.
; REGISTRATION NUMBER: 30,819
; REFERENCE/DOCKET NUMBER: 1017P021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)554-3300
; TELEFAX: (312)554-3301
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4456 bp
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-321-686B-1
```

Alignment Scores:

Pred. No.:	327	Length:	4456
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.08%	Indels:	0
DB:	3	Gaps:	0

US-09-784-810A-2 (1-384) x US-08-321-686B-1 (1-4456)

OY 231 GlnGlyProValAspAlaHisLeu 238
 |||||
 DB 2797 CAGGTCCTGTGTGAGCCTCACCTA 2820

RESULT 10
 US-09-735-934A-3/c
 ; Sequence 3, Application US/09735934A
 ; Patent No. 6372468
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, JiaYin et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: C1000851
 ; CURRENT APPLICATION NUMBER: US/09/735,934A
 ; CURRENT FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 43950
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-735-934A-3

Alignment Scores:

Pred. No.:	2.68e+03	Length:	43950
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.08%	Indels:	0
DB:	4	Gaps:	0

US-09-784-810A-2 (1-384) x US-09-735-934A-3 (1-43950)

OY 2 AspProAlaGlyGlyProArgGly 9
 |||||
 DB 1151 GACCGCGCGGGGCGCCCGGGGG 1128

RESULT 11
 US-09-345-882-1
 ; Sequence 1, Application US/09345882
 ; Patent No. 6399373
 ; GENERAL INFORMATION:
 ; APPLICANT: Bouguetelerc, Lydie
 ; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
 ; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
 ; FILE REFERENCE: GENSET.031A
 ; CURRENT APPLICATION NUMBER: US/09/345,882
 ; CURRENT FILING DATE: 1999-06-30
 ; PRIOR APPLICATION NUMBER: US 60/091,315
 ; PRIOR FILING DATE: 1998-06-30
 ; PRIOR APPLICATION NUMBER: US 60/111,909
 ; PRIOR FILING DATE: 1998-12-10
 ; NUMBER OF SEQ ID NOS: 140
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 1
 ; LENGTH: 162450
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: allele
 ; LOCATION: 72794

OTHER INFORMATION: 5-124-273 : polymorphic base A or G

FEATURE: NAME/KEY: allele
 LOCATION: 88073
 OTHER INFORMATION: 5-127-261 : polymorphic base A or C

FEATURE: NAME/KEY: allele
 LOCATION: 90842
 OTHER INFORMATION: 99-1437-325 : polymorphic base A or G

FEATURE: NAME/KEY: allele
 LOCATION: 93714
 OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT

FEATURE: NAME/KEY: allele
 LOCATION: 97122
 OTHER INFORMATION: 99-1442-224 : polymorphic base G or T

FEATURE: NAME/KEY: allele
 LOCATION: 97152
 OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T

FEATURE: NAME/KEY: allele
 LOCATION: 99098
 OTHER INFORMATION: 5-130-257 : polymorphic base A or G

FEATURE: NAME/KEY: allele
 LOCATION: 99117
 OTHER INFORMATION: 5-130-276 : polymorphic base A or G

FEATURE: NAME/KEY: allele
 LOCATION: 103806
 OTHER INFORMATION: 5-131-395 : polymorphic base A or T

FEATURE: NAME/KEY: allele
 LOCATION: 106940
 OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A

FEATURE: NAME/KEY: allele
 LOCATION: 108106
 OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A

FEATURE: NAME/KEY: allele
 LOCATION: 108149
 OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT

FEATURE: NAME/KEY: allele
 LOCATION: 108308
 OTHER INFORMATION: 5-135-357 : polymorphic base A or G

FEATURE: NAME/KEY: allele
 LOCATION: 108471
 OTHER INFORMATION: 5-136-174 : polymorphic base C or T

FEATURE: NAME/KEY: allele
 LOCATION: 134134
 OTHER INFORMATION: 5-140-120 : polymorphic base C or T

FEATURE: NAME/KEY: allele
 LOCATION: 134362
 OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A

FEATURE: NAME/KEY: allele
 LOCATION: 134374
 OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA

FEATURE: NAME/KEY: allele
 LOCATION: 146328
 OTHER INFORMATION: 5-143-84 : polymorphic base A or G

FEATURE: NAME/KEY: allele
 LOCATION: 146345
 OTHER INFORMATION: 5-143-101 : polymorphic base A or C

```

FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 9-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 9-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:

```

```

NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Alignment Scores:
Pred. No.: 8.84e+03 Length: 162450
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
DB: 4 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-345-882-1 (1-162450)
QY 64 ValArgSerGluIleuGlyArg 71
DB 272 GTCGGAGTACGAGCTCGTGC 295

RESULT 12
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FRISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103, 840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

Alignment Scores:

Pred. No.: 1.66e+05 Length: 4403765
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
DB: 4 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-103-840A-2 (1-4403765)

OY 7 ProArgGlyValLeuProArgPro 14

DB 430396 CCTCGCGGGGTTCTTCACGACCA 430373

RESULT 13

US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:

Pred. No.: 1.66e+05 Length: 4411529
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.08% Indels: 0
DB: 4 Gaps: 0

US-09-784-810A-2 (1-384) x US-09-103-840A-1 (1-4411529)

OY 7 ProArgGlyValLeuProArgPro 14

DB 430313 CCTCGCGGGGTTCTTCACGACCA 430290

RESULT 14

US-08-177-109A-3/c
Sequence 3, Application US/08177109A
Patent No. 5869615
GENERAL INFORMATION:
APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby
TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/177,109A

FILING DATE: 03-JAN-1994

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: WU 107

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 873-8794

TELEFAX: (404) 873-8795

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 33 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-177-109A-3

Alignment Scores:

Pred. No.: 35.2 Length: 33
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.82% Indels: 0
DB: 2 Gaps: 0

US-09-784-810A-2 (1-384) x US-08-177-109A-3 (1-33)

OY 208 GlyValGluAlaThrLeuPro 214

DB 26 GGTGACTGCGGACTCTTCCG 6

RESULT 15

US-08-687-706-3/c
Sequence 3, Application US/08687706
Patent No. 5928892
GENERAL INFORMATION:
APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby
TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,706
FILING DATE: 26-JUL-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/177,109
FILING DATE: 03-JAN-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: WU 107 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

Mon Jul 14 09:16:35 2003

us-09-784-810a-2.olig.rni

Page 8

LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-687-706-3

Alignment Scores:
Pred. No.: 35.2
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.82%
DB: 2
Length: 33
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-784-810A-2 (1-384) x US-08-687-706-3 (1-33)

QY 208 GYARGLEuAlArhLeuPro 214
|||||
Db 26 GGTGACCTCGGACCTTCG 6

Search completed: July 12, 2003, 21:48:22
Job time : 1510 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:17:53 : Search time 292 Seconds

(without alignments)
2961.532 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 384
Sequence: 1 MDPAGPGRVGLPRRCRVLYL.....CVPPEPSMKPQOMPPEEPL 384

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 segs, 112599159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4368727

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=rlh
-Q=/cgn1/USPto.spool/US09784810/runat_11072003_103504_27188/app_query.fasta.1.583
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=olig.rng -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09784810_@C@N_1_1_253_@runat_11072003_103504_27188 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEIO -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	384	100.0	1600	22 AAD14424	Human sphingosine
2	144	37.5	1173	24 ABL59532	Human sphingosine
3	144	37.5	1205	22 AAC84161	Human sphingosine
4	144	37.5	1573	21 AAT75676	DNA encoding a hum
5	144	37.5	1719	21 AAD04477	Human sphingosine
6	144	37.5	1722	22 AAH16415	Human cDNA sequenc
7	144	37.5	1821	22 AAH15652	Human cDNA sequenc
8	141	36.7	1562	24 ABL90618	Human polynucleoti
9	117	30.5	1438	22 ABL59336	Human polynucleoti
10	115	29.9	1200	22 AAT11122	Human polynucleoti
11	115	29.9	1533	20 AAV84490	Human secreted pro
12	115	29.9	1533	22 ABA83273	Human secreted pro
13	106	27.6	1447	21 ABA50508	Human sphingosine
14	80	20.8	240	22 AAD04478	Conserved region D
15	75	19.5	675	22 AAH07214	Human cDNA clone (
16	63	16.4	296	22 AAS24477	Human ovarian PCR-
17	63	16.4	296	22 AAB33083	Human ovarian tumo
18	56	14.6	587	22 AAH10443	Human cDNA clone (
19	43	11.2	604	22 AAB04943	Human cDNA clone (
20	39	10.2	394	24 ABL66446	Lung cancer relate
21	33	8.6	1559	21 AA247168	Mouse sphingosine
22	33	8.6	1759	22 AAD14425	Mouse consensus sp
23	33	8.6	1815	21 AA247167	Mouse sphingosine
24	19	4.9	30	24 ABA41630	Human SPHK1 cDNA c
25	9	2.3	65	24 ABA56403	Mouse spliced tran
26	9	2.3	1117	22 AAS02402	Human secreted pro
27	9	2.3	1280	21 AAC33647	Arabidopsis thalia
28	9	2.3	1394	24 ABO54324	Human ovarian anti
29	9	2.3	1521	22 AAD08644	Human kinase (PKIN
30	9	2.3	1578	23 AAS85330	DNA encoding novel
31	9	2.3	2136	22 AAK94172	Human full-length
32	9	2.3	2380	22 AAS14817	Human cDNA for sph
33	9	2.3	2380	24 ABL59533	Human sphingosine
34	9	2.3	2422	23 AAS85331	DNA encoding novel
35	9	2.3	5330	19 ABA59103	Bovine differentia
36	9	2.3	11096	22 ABA18278	Human nervous syst
37	9	2.3	18609	22 AAS21769	Human gene for col
38	9	2.3	25	21 AAS05512	Sphingosine kinase
39	8	2.1	25	21 AAS05513	Probe, SK BOX, for
40	8	2.1	28	22 AAD04490	Human SNP oligonuc
41	8	2.1	51	22 AAL30926	Sequence of intron
42	8	2.1	149	14 AAO43676	Sequence of human
43	8	2.1	242	9 AAN80878	Human secreted pro
44	8	2.1	244	21 AAC10623	
45	8	2.1	244	21 AAC10623	

ALIGNMENTS

RESULT 1	
AAD14424	
ID AAD14424 standard; cDNA: 1600 BP.	
AC AAD14424:	
XX 01-NOV-2001 (first entry)	
DT	
XX	
DE Human sphingosine kinase (Sphk) cDNA #1.	
XX	
XX Human; sphingosine kinase; Sphk; restenosis; ischemia; gene therapy;	
KW antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;	
KW cytosolic; tumor necrosis factor-alpha; TNF; atherosclerosis; lymphoma;	
KW leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.	
OS Homo sapiens.	
XX	
XX	
FH Key Location/Qualifiers	

FT CDS 176..1330
FT /+tag= a
FT /product= "human sphingosine kinase (Spk) protein #1"
XX WO200160990-A2.
XX 23-AUG-2001.
XX 14-FEB-2001; 2001WO-US04789.
XX 14-FEB-2000; 2000US-0182360.
XX 22-MAR-2000; 2000US-0191261.
XX (CURA-) CURAGEN CORP.
XX (GETH) GENENTECH INC.
XX Rastelli L;
XX WPI: 2001-514770/56.
XX P-PSDB; AAE07882.
XX
XX An isolated sphingosine kinase polypeptide useful for treating a
PT Spk-associated disorder especially cancer, restenosis or ischemia in a
PT human -
PS Claim 8; Page 89-90; 107pp; English.
XX
XX The present invention relates to sphingosine kinase (Spk) polypeptides
CC and nucleic acids encoding them. Spk is useful for treating a Spk-
CC associated disorder especially cancers such as leukemia, lymphoma,
CC ovarian, breast, lung, colon, testicular, stomach and skin.
CC atherosclerosis, restenosis or ischemia and cell proliferative disease
CC or disorder associated with vascular diseases. Spk gene is used in gene
CC therapy and antisense-therapy. Sphingolipids serving as signalling
CC molecules, have recently emerged as regulators of cell growth,
CC differentiation, diverse cell phenotypes and cell death. Activation of
CC Spk by tumor necrosis factor (TNF) alpha inhibits apoptosis in human
CC endothelial cells. The present sequence is human sphingosine kinase
CC (spk) cDNA.
XX
XX Sequence 1600 BP; 265 A; 492 C; 531 G; 311 T; 1 other;
SQ
Alignment Scores:
Pred. No.: 0 Length: 1600
Score: 384.00 Matches: 384
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB:
US-09-784-810a-2 (1-384) x AAD14424 (1-1600)
QY 1 MetaspProalaglyglyProargglyValLeuProargProCysArgValLeuValLeu 20
DB 176 ATGATCCAGCGGCGGCGCCCGGGGCTCTCCCGGCGCCGCGGCTGCGTGG 235
QY 21 LeuAnpProargglyglyLylysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
DB 236 CTGAACCCGCGGCGGCGGAGGCAAGGCTTCGAGCTTCGAGTCACTGCAGCC 295
QY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgATGASNHASAla 60
DB 296 CTTTGGCGTGAAGTCTCCCTTCAGCTGATGCTCACTGAGCGGCGGAGACACGCG 355
QY 61 ArgGluLeuValArgSerGluGluLeuGluArgTrpAspAlaLeuValMetSerGly 80
DB 356 CGGAGCTGCTGGCGGAGAGAGCTGGCCCTGGAGCGGTCTGGTGCATGTCTGGA 415
QY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
DB 416 GACGGGCTATGACAGAGGTGTAAGCGGCTCATGAGAGCGGCTGACTGGAGACGCC 475
QY 101 IleGlnLysProLeuCySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120

DB 476 ATCCAGAGCCCTCTGTAGCTCCACAGAGCTCTGGCAACGGCTGCGACTTCCTTG 535
QY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
DB 536 AACCATTAAGCTGGCTATGAGAGAGGTACCAATGAAGACCTCTGACCAACATGGACCTA 595
QY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
DB 596 TTGCTGTGGCGCGCGGTGCTGTACCATGAACCTGCTGTCTGACACAGGCTTCGGGG 655
QY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu 180
DB 656 CTGCGCTGCTTCTCTGTGCTCAGCCGTGGCGGCTTCAATGCTGATGTGGACCTAGAG 715
QY 181 SerAspLysTyrArgArgLeuGlyGlyMetArgPheThrLeuGlyThrPheLeuArgLeu 200
DB 716 AGTATGAATGATGCGGCTGTGGGAGATGCCCTTCACCTCTGGGACCTTCCTGCTGTG 775
QY 201 AlaAlaLeuValArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
DB 776 GCAGCCCTGGGACACCTACCGCGCGGAGTGGCTACCTCCCTGTAGGAAGAGTGGCTTC 835
QY 221 LysThrProAlaSerProValValGlnGlnGlyProValAspAlaHisLeuValPro 240
DB 836 AAGACACCTGCTCCCGGTGTGTGCTCAGACAGGCGCGGTAGATGCACACCTGGTGCCA 895
QY 241 LeuGluGluGlnValProSerHisTrpGlnValValProAspGluAspPheValLeuVal 260
DB 896 CTGGAGAGAGAGAGGCTTCTCACTGCGAGGTGTGCTCCAGAGAGACTTGTGTCTAGTC 955
QY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaPrometGlyArgCys 280
DB 956 CTGGACACTGCTGCACCTCCACCTGGCCAGTGAATGTTCTCTGACCCATGGGCGCTGT 1015
QY 281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu 300
DB 1016 GCAGCTGGGCTATGCATCTGTTCACGTGGCGGAGAGTGTCTGTGCTGATGCTGCTG 1075
QY 301 ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVal 320
DB 1076 CGCCTTCTCTGCGCATGGAGAGGCGAGCATATGAGATGAATGCCCTACTTGCTGA 1135
QY 321 TyrValProValValAlaPheArgLeuGluProLysAspGlyLyGlyValPheAlaVal 340
DB 1136 TATGTGCCCTGCTGCTGCTTCGCTTGGAGCCCAAGATGGAAAGTGTGTGTGCACTG 1195
QY 341 AspGlyLeuLeuMetValSerGluAlaValGlnGlnValHisProAsnTyrPheTrp 360
DB 1196 GATGGGGAATGTATGTATGACGAGCGCTGCGAGGCGCAGCTGCACCAACTACTCTGG 1255
QY 361 MetValSerGlyCysValGluProProSerTrpLysProGlnGlnMetProProPro 380
DB 1256 ATGGTACGGGTTGCGTGGAGCCCGCGCAGCTGGAAGCCCAACAGATGCGACCGCCA 1315
QY 381 GluGluProLeu 384
DB 1316 GAAGAGCCCTTA 1327
RESULT 2
ABL59532 standard; cDNA: 1173 BP.
ID ABL59532;
AC ABL59532;
XX 16-JUL-2002 (first entry)
XX Human sphingosine kinase (Spk1) cDNA SEQ ID NO:32.
XX Human sphingosine kinase (Spk1); enzyme: chromosome 17q25.2; gene:
XX tumour; lipid associated gene; lipid metabolism; lipid synthesis; ss.
OS Homo sapiens.


```

XX  MO200227028-A1.
XX  04-APR-2002.
XX  27-SEP-2001; 2001MO-US30366.
XX  28-SEP-2000; 2000US-0676052.
XX  (ATAI-) ATAIRGIN TECHNOLOGIES INC.
XX  Skinner MK, Patton JT, Chaudhary J;
XX  WPI; 2002-402054/43.
XX  WPI; 2002-402054/43.
XX  Identifying tumor characteristics in a tissue sample taken from a
XX  patient, involves determining the copy number or expression level of
XX  genes associated with lipid metabolism, synthesis or action
XX  Example 1; Page 90; 113pp; English.
XX  The present invention describes a method for identifying tumour
XX  characteristics, comprising measuring a copy number or expression level
XX  of at least two genes associated with lipid metabolism, synthesis, or
XX  action in cells from a patient tissue sample, and comparing the results
XX  with a copy number or expression level of the genes in a normal cell.
XX  Also described is an array of nucleic acid polymers immobilised on a
XX  solid support, comprising a solid support, at least two different nucleic
XX  acid polymers which are each specific for a different gene associated
XX  with lipid metabolism, synthesis or action, where each nucleic acid
XX  polymer is located at a predetermined position on the solid support, and
XX  the array comprises nucleic acid polymers which are specific for less
XX  than 100 genes other than the selected genes. The method is useful for
XX  determining tumour characteristics in a tissue sample taken from a
XX  patient. The present sequence represents a human lipid-associated gene
XX  related cDNA sequence, which is used in the exemplification of the
XX  present invention.
XX  Sequence 1173 BP; 188 A; 349 C; 390 G; 246 T; 0 other.
XX
XX  Alignment Scores:
XX  Pred. No.: 5,31e-134 Length: 1173
XX  Score: 144.00 Matches: 144
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 37.50% Indels: 0
XX  DB: 24 Gaps: 0
XX
US-09-784-810A-2 (1-384) x ABL59532 (1-1173)
OY 1 Metaspromalaglygylproarglyvalleuproargprocyarsargvalleuallenu 20
DB 19 ATGGATCCAGCGGGGCGGCCCGGGGGGCTGCTCCCGGCGCTCCCGGCTGCTGCTG 78
OY 21 Leuasnproargglyglylysglylysalaleuenuleupheargserhisvaligipro 40
DB 79 CTGAACCCCGCGCGCGGCAAGGCAAGCCCTTGACGCTCTCCGAGTCACAGTCAGCC 138
OY 41 Leuenuaagualagluileserphenrleumetleuthrgluaragarsnhsiala 60
DB 139 CTTTGGCTGAGCGCTGAATCTCTTACGCTGATGCTCACTGAGCGCGGCAACACGCG 198
OY 61 Arggluenuvalargsergluileuglyvargrtrpsalaleuvalalemetsergly 80
DB 199 CGGAGGCTGGTGGCTGGAGAGAGCTGGGCGCTGGAGCGCTGCTGGTGTATGTCTGGA 258
OY 81 Aspilyleumethisgluivalvalasnilyleumetgluarqproasptprglutthrala 100
DB 259 GACGGGCTGATGACAGAGGTGTGACAGCGGCTCATGACGCGGCTGACTGGACACCC 318
OY 101 Ileglulysproleucysserleuproalaglyserglyalanaalaleuallaserleu 120
DB 319 ATCCAGAGAGCCCTGTGTAGCTCCAGACAGGCTGTGCAACGCGGCTGGCAGCTTCCTTG 378

```

```

OY 121 AsnhistyrAlaglytyrGluInValThrAsnGluAspleuLeuThrAsnCysThrleu 140
DB 379 AACCATATGCTGCTATGACAGGTCACCAATGAGACCTCTGACCAACTGACGCTA 438
OY 141 Leuleucysarg 144
DB 439 TTGCTGTGCGCC 450
RESULT 3
AAC84161
ID AAC84161 standard; cDNA; 1205 BP.
XX
AC AAC84161;
XX
XX 19-MAR-2001 (first entry)
XX
XX Human sphingosine kinase encoding cDNA.
XX
XX Sphingosine kinase; antirheumatic; antiasthmatic; antiatherosclerotic;
XX antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
XX human; ss.
XX Homo sapiens.
XX
XX Key location/Qualifiers
XX CDS 33..1187
XX /*tag= a
XX /product= "sphingosine kinase"
XX
MO200070028-A1.
XX 23-NOV-2000.
XX
XX 12-MAY-2000; 2000MO-AU00457.
XX
XX 13-MAY-1999; 99AU-0000339.
XX 08-JUL-1999; 99AU-0001504.
XX
XX (JOHN J) JOHNSON & JOHNSON RES PTY LTD.
XX
XX Pitson SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vadas MA;
XX
XX WPI; 2001-016227/02.
XX P-PSDB: AAB48007.
XX
XX Novel sphingosine kinase protein and nucleic acid molecules for
XX diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,
XX atherosclerosis, inflammation, meningitis, multiple sclerosis and
XX septic shock
XX
XX Claim 4; Fig 7a; 100pp; English.
XX
XX This cDNA encodes a human sphingosine kinase (SK) protein. The human SK
XX protein, encoding nucleic acids and modulators are useful for modulating
XX expression, functional activity or cellular functional activity of
XX sphingosine kinase in a subject and also for treating a mammal by
XX modulating the activity of SK. Diseases treated by regulating SK
XX cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
XX inflammation, meningitis, multiple sclerosis and septic shock.
XX
XX Sequence 1205 BP; 196 A; 357 C; 400 G; 252 T; 0 other.
XX
XX Alignment Scores:
XX Pred. No.: 5,44e-134 Length: 1205
XX Score: 144.00 Matches: 144
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 37.50% Indels: 0
XX DB: 22 Gaps: 0
XX
US-09-784-810A-2 (1-384) x AAC84161 (1-1205)

```



```

QY 180 GluSerAsp-LysTyrArgArgLeuGluMetArgPheThrLeuGlyThrPheLeuAr
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 667 GAGAGTGA-GAATATACGGCGCTGGGGGATGCTTACTCTGGGCACTTCCGCG
QY 199 GluAlaAlaLeuArgTyrArgGlyArgLeuAlaThr-LeuProValGlyArgValG
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 726 TCTGGACGCCCTGCGACCTACCGGGCGGACATGGCCTA-CTCCCTGTAGGAAGAT-G
QY 219 LysPhe-LysThrProAlaSerProValValAlaGlnGlnGlyProValAlaPalaHsLeu
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 784 GGTTCCAGACACCTGCTCCGCCGTGTGTGTGTCAGAGAGGCCGGGTAGATGACACACTT
QY 239 ValProLeuGluGlnGln-ValProSerHisTrpGln-ValValProAspGlnAspPhe
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 844 GTCCCACTGAGAGAGCC-AGTGCCTCTCACTGAGAC-AGTGTGCCCCGACGAGAGACTTGG
QY 258 AluValLeuAlaLeuAlaLeuHisSerHisLeuAla-SerGluMetPheAlaAlaProMet
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 902 TGCTAGTCTGCGACCTGCTGCACTCGCACCTGGG-CAGTAGAGATGTTTGTGCACCCATG
QY 278 GlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAla
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 GCGCGCTGTGCACTGCGCATGCAATCTTCTACCTGCGGGGAGTGTCTCTGCGCC
QY 298 MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGlnCysPro
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 ATCTGCTGCGCTCTTCTCTGCGCATGAGAGAGGCGCATGTGAGTGAATGAGATGCGCC
QY 318 TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 TACTTGATATATGCGCCGTGCGCTTCCGCTTGAGGCCAAGATGGGAAGATGTG
QY 338 PheAlaValAspGlyGluLeuMetValSerGluAlaValAlaGlnGlyValHisProAsn
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 TTTCAGTGGATGGGAAATGATGATTGAGAGAGCCGTGAGGCGCCGAGTGCACCCCAAC
QY 358 TyrPheTyrMetValSerGlyCysValGluProProSerTrpLysProGlnGlnMet
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 TACTTGTGATGTGTCAGCGCTTGTGTCAGACCCCGCCCAAGCTGTGAAGCCCAAGATG
QY 378 ProProProGluGluProLeu 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1261 CCACCGCCAGAGAGAGCCCTTA 1281

RESULT 5
AAD04477
ID AAD04477 standard; cDNA; 1719 BP.
AC AAD04477;
XX
XX 04-JUL-2001 (first entry)
DT
XX
XX Human sphingosine kinase type 1 (hsk1) cDNA.
DE
XX
XX Human; sphingosine kinase type 1; sk1; chromosome 17q25.2;
KM sphingosine-1-phosphate; S1P; drug screening; therapy; haemostasis;
KM thrombosis; allergic reaction; proliferative disease; cancer;
KM haematopoietic disorder; leukaemia; cardiovascular disease; stroke;
KM atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
KM autoimmune disease; inflammatory disease; multiple sclerosis;
KM T helper-1 related disease; chronic obstructive pulmonary disease;
KM asthma; myocardial infarction; neurodegenerative disorder;
KM wound healing; embryogenesis; anticoagulant; cerebroprotective;
KM neuroprotective; antiparietal; antiarthritic; cyostatic; cardiant;
KM vulnerable; ss.
XX
XX Homo sapiens.
XX
XX
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 270..1424
FT /*tag= a
FT /*product= "Human sphingosine kinase type 1 (hsk1)"

```

```

FT FT /note= "CDS is specifically claimed in claim 2 and
FT FT shown as SEQ ID NO 2"
FT FT polyA_signal
FT FT 1675..1681
FT FT /*tag= b
FT FT misc_feature
FT FT 264..273
FT FT /*tag= c
FT FT /note= "Translational Initiator ATG is in a partial
FT FT Korak consensus"
XX
XX W0200131029-A2.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000MO-EP09498.
XX
XX 28-OCT-1999; 99US-0162307.
XX 07-FEB-2000; 2000US-0180525.
XX
XX (WARN ) WARNER LAMBERT CO.
XX
XX Allen J, Gosink M, Melendez AJ, Takacs L;
XX WPI; 2001-300510/31.
XX DR P-PSDB; AAE00924.
XX
XX New human sphingosine kinase type I gene for screening drug candidates
XX particularly inhibitors used for preventing or treating e.g.
XX PT atherosclerosis, thrombosis, asthma and diabetes
XX
XX Claim 2; Fig 1; 91pp; English.
XX
XX The present sequence is human sphingosine kinase type 1 (hsk1) cDNA.
XX The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the
XX substrate sphingosine to sphingosine-1-phosphate (S1P). The sk1 gene
XX and encoded polypeptide are applicable in screening drug candidates
XX particularly inhibitors for preventing or treating disorders such as
XX haemostasis, thrombosis, allergic reactions, proliferative diseases
XX including cancer, haematopoietic disorders such as leukaemia,
XX cardiovascular diseases such as stroke, atherosclerosis and coronary
XX artery disease, dyslipidaemia, diabetes including type I and type II
XX diabetes, autoimmune and inflammatory diseases such as multiple
XX sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
XX disease, asthma, myocardial infarction, neurodegenerative disorders,
XX natural wound healing processes and embryogenesis.
XX
XX SQ Sequence 1719 BP; 305 A; 529 C; 556 G; 329 T; 0 other;

Alignment Scores:
Pred. No.: 7.59e-134 Length: 1719
Score: 144.00 Matches: 375
Percent Similarity: 95.428 Conservative: 0
Best Local Similarity: 95.428 Mismatches: 9
Query Match: 37.508 Indels: 18
DB: 22 Gaps: 0

US-09-784-810A-2 (1-384) x AAD04477 (1-1719)
QY 1 MetAspProAlaGlyGlyProAlaGlyValLeuProArgProCysArgValLeuValLeu 20
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 270 ATGATATCACCGGCGGCCCGCGGGCGTGTCCCGGCGCTCCCGCTGTGCTG
QY 21 LeuAsnProArgGlyGlyLysGlyValAlaLeuGlnLeuPheArgSerHisValGlnPro 40
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 330 CTGAACCCGCGCGCGGCGGCAAGGAGCCCTTGCAAGCTCTTCCGAGTCAAGTCCAGCCC 389
QY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgHisAla 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 390 CTTTGGCTGAGCGTGAATCTCTTACAGCTGATGTCTCACTAGAGCGGGAACACAGCGG 449
QY 61 ArgGluLeuValAlaArgSerGluGluLeuGlyArgTyrPaspAlaLeuValMetSerGly 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 450 CGGAGACTGTGCGGTGAGAGAGACTGCGCGCGTGGAGAGCGCTGTGATCTGTGGA 509

```

QY 81 AspGlyLeuMethisGluValValAlaSerGlyLeuMetGluProAspProGlnThrAla 100
 DB 510 GACGGCTGATGACACAGGTGTGAGACGGCTCATGAGCGCTACGAGGAGACCGCC 569
 QY 101 IleGlnLysProLeucysSerLeuProAlaGlySerGlyAlaAlaLeuAlaSerLeu 120
 DB 570 ATCCAGAGAGCCCTGTGTAGCTCCACAGGCTCTGTGGCAACGCTGGCACCTCCCTG 629
 QY 121 AsnHisTyrAlaGlyTyrGlnValThrAsnGluAspLeuThrAsnCysThrLeu 140
 DB 630 AACCATATGCTGCTGATGACAGGTGCACATATATATATATATATATATATATATAT 689
 QY 141 LeuLeucysArgProVal - LeuSerProMetAlaLeuLeuSerLeuHisThrAlaSerG 160
 DB 690 TTGCTGTGCGC - CCGGC - TGCTGTACCATGAGACCTGCTGTCTGTGACACAGGCTTGG 747
 QY 160 LysLeuArgSer - PheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeu 179
 DB 748 GCGTGGCGCCT - CTTCTCTGTGCTACGCTGGCGCTTCATCTCTGATGTGACCTTA 806
 QY 180 GluSerAsp - LysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuAr 199
 DB 807 GAGAGTGA - GAGATATGCGCGCTGTGGGAGATGCGCTTCTGCTGTGGCACCTTCTCGG 865
 QY 199 GLeuAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThr - LeuProValGlyArgValG 219
 DB 866 TCTGCGACGCTGCGACCTACCGCGCGCGACTGGCCTTA - CCTCCCTGTAGGAGAGT - G 923
 QY 219 LysPhe - LysThrProAlaSerProValValAlaGlnGlnGlnLysProValAspAlaHisLeu 238
 DB 924 GGTTCACAGACACTCTCCCTCCCTGTGTGCTCCAGAGGCGCGCTGATGATGACACCTT 983
 QY 239 ValProLeuGlnGln - ValProSerHisTyrGln - ValValProAspGlyAspPheV 258
 DB 984 GTGCGCATGAGAGAGGCC - AGTGCCCTCTCACTGAGAC - AGTGCGCCCGCAGAGAGACTTGG 1041
 QY 258 alleuValleuAlaLeuLeuHisSerHisLeuAla - SerGluMetPheAlaAlaProMet 277
 DB 1042 TGCTAGTCTGCGACGCTGCGACCTGCGACCTGGG - CAGTGAAGTGTTCCTGACCCCTG 1100
 QY 278 GlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAla 297
 DB 1101 GCGCGCTGTCACCTGGCTGCTATGATCTTCTACCTGCGCGCGGAGTGTCTGTGCC 1160
 QY 298 MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysPro 317
 DB 1161 ATGCTGCTGCGCTCTCTCTGCGCATGAGAGAGGCGCATATGAGATGAAATGCCCC 1220
 QY 318 TyrLeuValTyrValProValAlaAlaPheArgLeuGlnProLysAspGlyLysGlyVal 337
 DB 1221 TACTTGTATATGTGCGCTGTGCTGCCCTTCCGCTTGAGCCCAAGGATGGAGAGGTGTG 1280
 QY 338 PheAlaValAspGlyGluLeuMetValSerGluAlaValGlnGlnValHisProAsn 357
 DB 1281 TTTCAGTGGATGGGAGATATGATGTATGCGAGAGCCGTGAGGGCCAGGTGACCCCAAC 1340
 QY 358 TyrPheTyrMetValSerGlyCysValGlnProProSerTyrLysProGlnLysMet 377
 DB 1341 TACTTCTGATGTGTCAGCGGTGCTGAGAGCCCGCCAGCTGAGAGAGCCAGAGATG 1400
 QY 378 ProProGlnGlnProLeu 384
 DB 1401 CCACCGCCAGAGAGAGCCCTTA 1421
 RESULT 6
 AAH16415
 ID AAH16415 standard; cDNA; 1772 BP.
 AC AAH16415:
 DT 26-JUN-2001 (first entry)
 TX Human cDNA sequence SBO ID NO:15393.
 DE

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KM Homo sapiens.
 OS Homo sapiens.
 PN EP1074617-A2.
 PD 07-FEB-2001.
 PF 28-JUL-2000; 2000BP-0116126.
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 PA (HELI-) HELIX RES INST.
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 XX
 PS Claim 8; SEQ ID 15393; 2537bp + CD ROW; English.
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to 5'-end
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SO Sequence 1772 BP; 285 A; 566 C; 585 G; 336 T; 0 other:
 Alignment Scores:
 Pred. No.: 7, 81e-134 Length: 1772
 Score: 144.00 Matches: 375
 Percent Similarity: 95.42% Conservative: 0
 Best Local Similarity: 95.42% Mismatches: 9
 Query Match: 37.50% Indels: 18
 DB: 22 Gaps: 0
 US-09-784-810a-2 (1-384) x AAH16415 (1-1772)
 QY 1 MetAspProAlaGlyGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
 DB 350 ATGATATCCAGCGCGCGCGCGCGCTGCTCCCGCGCGCGCTGCTGCTGCTGCTG 409
 QY 21 LeuAsnProArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 40
 DB 410 CTGAACCCGCGCGCGCGCGCGCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTG 469

Oy	41	leuenuaiaagiuaiaagiulesePhehriueuethleuPrgiuaaArgAsnHiala	60
Db	470	CTTTGGCTAGAGCTGAATATCTCTTACAGCTGAATGCTCACTGAGAGCGGGAAACACAGCG	529
Oy	61	ArgsluIueuValaArgsergiuIuIueuIdyArgrTpaapalaleuValaImetSerily	80
Db	530	CGGAGAGCTGGCGGCTGGAGAGACTGGGCGCTGGAGACGCTGTGTGTATGTCATG	589
Oy	81	AspIyIueuMethIsuIuaValaAsnIyIueuMetGuaaGProaPrrgIuThraIa	100
Db	590	GACGGGCTGAGACAGAGGTGGTAAACGGGTCTATGGAGCGGCTGATGGGAAACACGCC	649
Oy	101	IleGIuIusProIeucYserIeupProIaIySerGIaAsnaIaleuAlaIaSerIeu	120
Db	650	ATCCAGAAAGCCCTGTGTAGCTCCACAGAGCTCTGGCAACGGGCTGGACGCTTCCTTG	709
Oy	121	AsnHsIryaIaagiIyTyrGIuIuaValThraAsnIuAspIeueThraAsnCysThreIu	140
Db	710	AACCATATATGTGGCTATAGACAGATCACATATAGACCTCTGACCACATGCACGCTA	769
Oy	141	LeuIeucYsaArgProVal--IeuserProMetAnIeuleuSerIeuHsIhrrAlaSerG	160
Db	770	TTGGCTGTGGCG-CCGGC-TGCTGTACCCCAAGAACCTGTCTCTCTGACACGGCTTCGG	827
Oy	160	IyIeuaArgSer-PheSerValIeuserIeuaIatrgIyPheIleAlaAspValaIspIeu	179
Db	828	GGCGTGGCGCT-CTTCTCTGTCTGACGCTGGCGCTGGGGCTTCATTTGCGATGTGGACCTA	886
Oy	180	GIuserAsp-LysTyrIarGarIeueIyGIuMetarGpHeThreueIyThrPheIueAr	199
Db	887	GAGAGTGA-GAAGATACGGCGCTGGGGGAATCGCTTCACTCTGGGACACTTCCTGGC	945
Oy	199	glEuaIaIaIeuaIrgrThrTyrArgGIyArgIeuaIaIathr-IeuProValGIyArGValG	219
Db	946	TCGTGGACACCTGGGCACACTACCGCGGCACACTGGCTTA-CTCTCCTGTAGGAAGAGT-G	1003
Oy	219	IyPhe-IySthrrProIaIaSerProValaValaGIuIuGIuIyProValaIspAlaHsIeu	238
Db	1004	GGTTCCAAAGACACTGCTCCCGCTGTGTGTCTCCAGCAGGGCCCGGTGATGACACTT	1063
Oy	239	ValProIeugIuIuGIuIu-ValProSerHsITrgIuIu-ValValProaPrrgIuIuAspPheY	258
Db	1064	GTGCCACACTGGAGGAGACC-AGTGCCCTCTCACTGGAC-AGTGATGCCAGACAGGACTTGG	1121
Oy	258	aIeuaValIeuaIaIeueuHsIserHsIleuAla-SerGIuMetPheAlaIaIaPromet	277
Db	1122	TGCTAGTCTGTGGCACTGCTGCACCTGCACCTGGG-CAGTGAGATGTTTGTGTGACCCATG	1180
Oy	278	GIyArGcySaIaIaagiIyValaMethIsIeueheryIyValaIaagiIyValaSerIyGaIa	297
Db	1181	GGCGGCTGTGACGTGGCGGTCAATCTGTTCTAACGTCGGGGGGGAGTCTTCGTCGCC	1240
Oy	298	MetIeueuAqIeupheIueuAlaMetGIuIySgIyArGHaIeMetGIuTyrGIuIcYsPro	317
Db	1241	ATGCTGCTGGCGCTCTTCTCTGCGCATGAGAAAGGACAGCATATAGATATGAATGCCCC	1300
Oy	318	TyrIeuaValIyTyrValProValaValaIaPheArIeueGIuIuProIyAspGIyIySgIyVala	337
Db	1301	TACTTGTGTATACGTGCCCTGGTGGTGGCTTCGCTTGGAGCCCAAGAGATGGGAAAGGTGTG	1360
Oy	338	PheIaIaValaAspGIyGIuIeueMetValserGIuIaIaValaGIuIuGIuIyValaHsIProaSn	357
Db	1361	TTTGGCAGTGGATGGGGAATTATGTATTAGCAAGGCGGTGCAGGGGCCAAGTCAACCCAAAC	1420
Oy	358	TyrPheTrrMetValSerGIyCysValaGIuIuProProaSerTrrPlySPrGIuIuIuMet	377
Db	1421	TACTTCTGTAGATGATACGGGTGGTGGTGGAGCCCGCCACACTGGAAGCCCCAGACATG	1480
Oy	378	ProProProGIuIuIuProIeu	384
Db	1481	CCACCGCCAGAGAGCCCTTA	1501

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukemias and CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed CC specification.

XX Sequence 1438 BP; 251 A; 428 C; 463 G; 296 T; 0 other;

Alignment Scores:

Pred. No.:	6,22e-107	Length:	1438
Score:	117.00	Matches:	374
Percent Similarity:	94.92%	Conservative:	0
Best Local Similarity:	94.92%	Mismatches:	10
Query Match:	30.47%	Indels:	20
DB:	22	Gaps:	0

US-09-784-810A-2 (1-384) x AA159336 (1-1438)

```

QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
   |||||
DB 1 ATGGATCCAGCGGGGCGCCCGCGGGGCTCTCCCGCGCCCTGGCCGCTGCTGCTG 60
   |||||
QY 21 LeuAsnProArgGlyGlyGlyValLeuGlnLeuPheArgSerHisValGlnPro 40
   |||||
DB 61 CTGAACCCCGCGGGGCGGAGGCGACGCTTCGACCTCTCCGAGTCAGTCACACCC 120
   |||||
QY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
   |||||
DB 121 CTTTGGGCTGAGGCTGAATCTCTTCACGCTGATGCTCATGACGCGGAGACACGCG 180
   |||||
QY 61 ArgGluLeuValArgSerGlyGlyLeuGluArgTyrPheAlaLeuValValMetSerGly 80
   |||||
DB 181 CGGGAGCTGCTGCTGGTGGAGAGCTGGGCGCTGGGAGAGCTGCTGCTATGCTCGGA 240
   |||||
QY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTyrGlnThrAla 100
   |||||
DB 241 GAGGGCTATACAGAGAGGTGTAGACGGCTATAGACGGCTGAGCTGAGAGAGCC 300
   |||||
QY 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAla-SerLe 120
   |||||
DB 301 ATCCAGAACCCCTCTGTAGCCCTCCAGAGGCTCTGGGACGCGCTGGCACCC-TCCTT 359
   |||||
QY 120 uAsnHisTyrAlaGlyTyrGlyGlnValThrAsnGluAspLeuThrAsnCysThrLe 140
   |||||
DB 360 GAACCATATATGCTGTAGAGCAGTGCACCAATGAAGACCTCTGACCACTGACAGCT 419
   |||||
QY 140 uLeuLeuCysArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSer 159
   |||||
DB 420 ATTGCTGTGCGG-CCGGC-TGCTGTACCCATGACCTGCTGTCTGTGCACACGGCTTCG 477
   |||||
QY 160 GlyLeuArgSer-PheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLe 179
   |||||
DB 478 GGGCTCGGCGCT-CTTCTGTGTGCTACACCTGGGCTTCATTTGCTGAGATGGAGCT 536
   |||||
QY 179 uGluSerAsp-LysTyrArgArgLeuGlyLumetArgPheThrLeuGlyThrPheLeu 199
   |||||
DB 537 AGAGAGTTGA-GAAGTATCGGGCTCGGGGGAGATGGCTTCACTGGGACACTTCTCTGC 595
   |||||
QY 199 rGluLeuAlaLeuLeuArgThrArgTyrArgGlyArgLeuAlaThr-LeuProAlaGlyArgVal 218
   |||||
DB 596 GTCTGGAGCCCTGGGACCTACCGGGCGGAGCTGAGCTTA-CTCTCTGTAGAGAGAT- 653
   |||||
QY 219 GlyPhe-LysThrProAlaSerProValValGlnGlnGlyProValAspAlaHisLe 238
   |||||
DB 654 GGGTTCCAAGACACTGCTCCCGCTGTGTGCTCCAGCGGGCGCGGTGAGATGCACACCT 713
   |||||
QY 238 uValProLeuGluGluGlnValProSerHisTyrGlnValValProAspGluAspPhe 257
   |||||
DB 714 TGTGCCACTGTGAGAGGCC-AGTGGCCCTCTCACTGAGC-AGTGTGCGCCAGAGACTTT 771
   |||||
QY 258 ValLeuValLeuAlaLeuLeuHisSerHisLeuAla-SerGluMetPheAlaAlaProMe 277
   |||||
DB 772 GTGTAGTCTGGACACTGTGCTGACTGCACCTGGG-CAGTGAATGTTTGTCTGCACCAT 830

```

```

QY 277 tGlyArgCysAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAl 297
   |||||
DB 831 GGGCCGCTGTGACAGTGGCGGTGATGATGTTCTACGTGCGGGGCGGAGTGTCTGCTGC 890
   |||||
QY 297 aMetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysPr 317
   |||||
DB 891 CATGCTGCTGGCGCTCTCTCTGCGCATGGAGAGGCGAGCATATGATGATGAAATGCC 950
   |||||
QY 317 cTyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
   |||||
DB 951 CTACTGTGATATGTGTGCGCGCTGCTGCGCTTCCGCTTGGAGCCCAAGATGGGAAAGTGT 1010
   |||||
QY 337 lPheAlaValAspGlyLysLeuMetValSerGluAlaValGlnGlyGlnValHisProAs 357
   |||||
DB 1011 GTTTGCAGTGGATGGGAAATGATGTAGCGAGGCGGTGCAGGGCCAGGTGCACCCAAA 1070
   |||||
QY 357 nTyrPheThrMetValSerGlyCysValGluProProSerTyrLysProGlnGlnMe 377
   |||||
DB 1071 CTACTCTGTGATGTGACGCGTTCGTGAGAGCCCGCCACAGTGAAGCCCGACAGAT 1130
   |||||
QY 377 tProProProGluGluProLeu 384
   |||||
DB 1131 GCCACCGCCAGAAAGCCCTTA 1152
   |||||
RESULT 10
AA161122 standard; cDNA; 1200 BP.
ID AA161122
AC AA161122;
DT 22-Oct-2001 (first entry)
DE Human polynucleotide SEQ ID NO 5111.
XX
XX Human, noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-Jul-2001.
XX
XX 26-Dec-2000; 2000MO-US34263.
XX
XX 21-Jan-2000; 2000US-0489725.
XX 25-Apr-2000; 2000US-0552317.
XX 09-Jul-2000; 2000US-0598042.
XX 19-Jul-2000; 2000US-0620312.
XX 03-Aug-2000; 2000US-0653450.
XX 14-Sep-2000; 2000US-0662191.
XX 19-Oct-2000; 2000US-0693036.
XX 29-Nov-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB: AAM41966.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 5111; 10078bp; English.
XX

```


CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1200 BP; 218 A; 353 C; 375 G; 254 T; 0 other;
Alignment Scores:
Pred. No.: 5.24e-105 Length: 1200
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.95% Indels: 0
DB: 22 Gaps: 0
US-09-784-810A-2 (1-384) x AA161122 (1-1200)
QY 270 SerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGlyValMetHisLeuPheTyr 289
Db 573 AGTACAGATGTTGTGTGACCCATGGCGCGTGTGACCTGCGCATGACATCTTCTAC 632
QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGlyLysGly 309
Db 633 GTGGGGCGGGAGGTCTCGTCCATGCTGCTGCGCTCTCCGCGCATGAGAGAGGC 692
QY 310 ArgHisMetGluTyrGluCysProTyrLeuValTyrValProValAlaAlaPheArgLeu 329
Db 693 AGGCATATGGAGTATGAAATGCCCTACTGTGTATACGTGCCCGGTGCGCTTCGCTTG 752
QY 330 GluProLysAspGlyLysGlyValAlaPheAlaValAspGlyLeuMetValSerGluAla 349
Db 753 GAGCCCAAGATGGTGAAGGTGTGTTCAGTGTGATGGGAAATGATGATGAGGAGGCC 812
QY 350 ValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro 369
Db 813 GTGAGAGGCCAGGTGACCCAAACTACTTCTGATGATGTCAGCGGTGCGTGAAGCCCGC 872
QY 370 ProSerTrpLysProGlnGlnMetProProProGluGluProLeu 384
Db 873 CCCAGCTGAGAGCCCGACAGATGCCACGCCAGAGAGGCCCTTA 917
RESULT 11
AAV84490
ID AAV84490 standard; DNA; 1533 BP.
XX
XX AAV84490:
XX
XX 01-MAR-1999 (first entry)
XX
XX Human secreted protein gene 80 clone HNP454.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.

XX
PN W09854963-A2.
XX
PD 10-DEC-1998.
XX
PF 04-JUN-1998; 98WO-US11422.
XX
PR 18-DEC-1997; 97US-0070923.
PR 06-JUN-1997; 97US-0048877.
PR 06-JUN-1997; 97US-0048881.
PR 06-JUN-1997; 97US-0048884.
PR 06-JUN-1997; 97US-0048893.
PR 06-JUN-1997; 97US-0048896.
PR 06-JUN-1997; 97US-0048899.
PR 06-JUN-1997; 97US-0048915.
PR 06-JUN-1997; 97US-0048949.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048972.
PR 06-JUN-1997; 97US-0049020.
PR 06-JUN-1997; 97US-0049375.
PR 05-SEP-1997; 97US-0057628.
PR 05-SEP-1997; 97US-0057635.
PR 05-SEP-1997; 97US-0057644.
PR 05-SEP-1997; 97US-0057647.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057661.
PR 05-SEP-1997; 97US-0057667.
PR 05-SEP-1997; 97US-0057761.
PR 05-SEP-1997; 97US-0057764.
PR 05-SEP-1997; 97US-0057770.
PR 05-SEP-1997; 97US-0057775.
PR 05-SEP-1997; 97US-0057778.
PR 06-JUN-1997; 97US-0048875.
PR 06-JUN-1997; 97US-0048878.
PR 06-JUN-1997; 97US-0048882.
PR 06-JUN-1997; 97US-0048885.
PR 06-JUN-1997; 97US-0048889.
PR 06-JUN-1997; 97US-0048894.
PR 06-JUN-1997; 97US-0048897.
PR 06-JUN-1997; 97US-0048900.
PR 06-JUN-1997; 97US-0048910.
PR 06-JUN-1997; 97US-0048962.
PR 06-JUN-1997; 97US-0048970.
PR 06-JUN-1997; 97US-0048974.
PR 06-JUN-1997; 97US-0049373.
PR 05-SEP-1997; 97US-0057584.
PR 05-SEP-1997; 97US-0057629.
PR 05-SEP-1997; 97US-0057642.
PR 05-SEP-1997; 97US-0057645.
PR 05-SEP-1997; 97US-0057648.
PR 05-SEP-1997; 97US-0057651.
PR 05-SEP-1997; 97US-0057662.
PR 05-SEP-1997; 97US-0057668.
PR 05-SEP-1997; 97US-0057762.
PR 05-SEP-1997; 97US-0057765.
PR 05-SEP-1997; 97US-0057771.
PR 05-SEP-1997; 97US-0057776.
PR 06-JUN-1997; 97US-0048876.
PR 06-JUN-1997; 97US-0048880.
PR 06-JUN-1997; 97US-0048883.
PR 06-JUN-1997; 97US-0048892.
PR 06-JUN-1997; 97US-0048895.
PR 06-JUN-1997; 97US-0048898.
PR 06-JUN-1997; 97US-0048901.
PR 06-JUN-1997; 97US-0048917.
PR 06-JUN-1997; 97US-0048963.
PR 06-JUN-1997; 97US-0048971.
PR 06-JUN-1997; 97US-0049019.
PR 06-JUN-1997; 97US-0049374.
PR 05-SEP-1997; 97US-0057627.
PR 05-SEP-1997; 97US-0057634.
PR 05-SEP-1997; 97US-0057643.
PR 05-SEP-1997; 97US-0057646.
PR 05-SEP-1997; 97US-0057649.

PR 05-SEP-1997; 97US-0057654.
PR 05-SEP-1997; 97US-0057666.
PR 05-SEP-1997; 97US-0057760.
PR 05-SEP-1997; 97US-0057763.
PR 05-SEP-1997; 97US-0057769.
PR 05-SEP-1997; 97US-0057774.
PR 05-SEP-1997; 97US-0057777.
PR (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA;
PI Fan P, Ferrle AM, Fischer CL, Florence CJ;
PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW;
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
DR WPI: 1999-059865/05.
DR P-PSDB: AAM8613.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 4: Page 343; 772pp: English.
XX
XX The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
CC encoding human secreted proteins (AAM88534 to AAM88756). The secreted
CC protein gene sequences are deposited with the ATCC under deposit numbers
CC ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, the
CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents a gene encoding a human secreted protein
CC (see descriptor line for gene number and clone identification).
XX
XX Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other:
SQ
Alignment Scores:
Pred. No.: 6,59e-105 Length: 1533
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29,95% Indels: 0
DB: 20 Gaps: 0
US-09-784-810a-2 (1-384) x AAV84490 (1-1533)
QY 270 SerGluMetPheAlaAlaPheMetGlyArgCysAlaAlaGlyValMetHisLeuPheTyr 289
DB 881 AGTGAATGTTTCTGCGCCACCATGCGCGCTGTCAGCGCTCATGATCTGTTCTAC 940
QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGly 309
DB 941 GTGCGGCGGAGATGCTCTGTCATGCTGCTGCGCTCTCTCTGCGCATGAGAGAGGC 1000
QY 310 ArgHisMetGluTyrGluCysProTyrLeuValTyrValProValAlaAlaPheArgLeu 329
|||||

DB 1001 AGGCATATGAGATGAATGCCCTACTGTGTATATGTGCCCTGTCGCTTCGCTTG 1060
QY 330 GluProLysAspGlyLysGlyValAlaPheAlaValAspGlyLeuMetValSerGluAla 349
DB 1061 GAGCCAGAGATGGAGAAAGGTGTGTGTGACATGATGGGAATGTATGTTGACGAGCC 1120
QY 350 ValGluGlyGlnValHisProAsnTyrPheTyrMetValSerGlyCysValGluProPro 369
DB 1121 GTGAGAGGCCAGGTGACCCCAACTCTCTGATGTGTGACGCTGTGAGAGCCCGC 1180
QY 370 ProSerTyrPheProGluGlnMetPheProProGluGluProleu 384
DB 1181 CCCAGCTGGAAGCCCGACAGATGACACCGCCAGAACAGCCCTTA 1225
RESULT 12
ID ABA83273
AC ABA83273 standard; cDNA; 1533 BP.
XX
XX ABA83273;
XX 07-FEB-2002 (first entry)
DE Human secreted protein gene 80 SEQ ID NO:90.
XX
XX Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
KW dermatological; immunosuppressive; antiinflammatory; immunostimulant;
KW cyostatic; cardiant; anti-vascular; anti-angiogenic; ophthalmological;
KW neuroprotective; nootropic; anticonvulsant; antialzheimer's; vulnerary;
KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW Gaucher's disease; cardiovascular disease; Schmitz syndrome; chemotaxis;
KW Chaga's cardiomyopathy; coronary arteriosclerosis; angiodysplasia;
KW corneal graft; neovascularization; diabetic retinopathy; degeneration;
KW neurological disorder; Huntington's chorea; Alzheimer's disease;
KW Parkinson's disease; infectious disease; chromosome 17; ss.
XX
XX Homo sapiens.
OS
XX WO200162891-A2.
XX 30-AUG-2001.
XX 21-FEB-2001; 2001WO-US05614.
XX 24-FEB-2000; 2000US-184836P.
XX 29-MAR-2000; 2000US-193170P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Ebner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;
PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrle AM, Fan P;
PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;
PI Zeng Z, Greene JM;
XX WPI: 2001-625724/72.
XX P-PSDB: ABB50380.
PT Nucleic acids encoding 207 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating, e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
PS
PS Claim 1: Page 965; 1533pp: English.
PS ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
CC activities based on the tissues and cells the genes are expressed in.
CC Example of these activities include: immunomodulatory; antisclerotic;
CC dermatological; immunosuppressive; antiinflammatory; immunostimulant;
CC anti-HIV; cyostatic; cardiant; anti-angiogenic; ophthalmological;
CC neuroprotective; nootropic; anticonvulsant; antialzheimer's; vascular;
CC antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used

CC in gene therapy and vaccine production. (I) and (II) can be used in the
CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC Gaucher's disease), cardiovascular diseases (e.g. Schmitz syndrome,
CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
CC disorders (e.g. corneal graft neovascularisation and diabetic
CC retinopathy), neurological disorders (e.g. Huntington's chorea,
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
CC ABA83193 and ABA850300 represent sequences used in the exemplification of
CC the present invention.

CC
XX
SQ Sequence 1533 BP: 272 A; 454 C; 495 G; 302 T; 10 other;

Alignment Scores:

Pred. No.:	6,59e-105	Length:	1533
Score:	115.00	Matches:	115
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	29.95%	Indels:	0
DB:	22	Gaps:	0

US-09-784-810A-2 (1-384) x ABA83273 (1-1533)

OY 270 SerGluMetPheAlaAlaProMetGlyArgGlyCysAlaAlaGlyValMetHisLeuPheTyr 289
DB 881 AGTGAATGTTTGGCTGCACCCATGGGCGCTGTGACGCTGGCGTCATGCTGTTCTTC 940
OY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGly 309
DB 941 GTGGCGGCGGAGTGTCTGTCGTCATGCTGCTGCCCTCTTCCTGCGCATGAGAAAGGC 1000
OY 310 ArgHisMetGluTyrGluCysProTyrLeuValTyrValProValAlaAlaPheArgLeu 329
DB 1001 AGGCATATGGAATGAATGAGCCCTTACTGTGATATGTCGCCGTGCTGCCCTTCGCTTG 1060
OY 330 GluProLysAspGlyLysGlyValPheAlaValAspGlyLysLeuMetValSerGluAla 349
DB 1061 GAGCCCAAGATGGAGAAAGGTGTCTTTCAGTGGATGGAGATTGATGTCAGAGGCC 1120
OY 350 ValGlnGlyGlnValHisProAsnTyrPheTyrMetValSerGlyCysValGluProPro 369
DB 1121 GTGCAGGCGCAGGTGCACCCAAATCTTCTGTGATGTGTCAGCGGTGCGTGAGCCCGC 1180
OY 370 ProSerTyrPysProGlnGlnMetProProProGluGluProLeu 384
DB 1181 CCCAGCTGGAACCCCGACAGATGCCACCGCCAGAAAGCCCTTA 1225

RESULT 13

AAA50508
ID AAA50508 standard; cDNA; 1447 BP.

AC AAA50508;

DT 05-DEC-2000 (first entry)

DE Human sphingosine kinase A cDNA.

XX Sphingosine kinase A; SKA; human; drug screening; infection;
XX antiinflammatory; antiallergic; anticancer; inflammation; allergy;
KW cancer; therapy; diagnosis; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 7..1161

PN WO200052173-A2.

XX 08-SEP-2000.

PF 02-MAR-2000; 2000WO-CA00223.

PR 02-MAR-1999; 99US-0122516.

PA (ALIX) NPS ALLELIX CORP.

PI Munroe D, Gupta A, Falzone GR;

DR MPI: 2000-572185/53.

DR P-PSDB: AAY06057.

PT New human sphingosine kinase A, B and C polynucleotides and
PT polypeptides useful in e.g. chromosome and gene mapping, and detecting
PT inflammation or disease associated with abnormal levels of sphingosine
PT kinase expression

XX Disclosure: Fig 1; 81pp; English.

XX The present sequence is that of an isolated polynucleotide encoding
XX human sphingosine kinase A (SKA, see AAY96057), an enzyme that
XX phosphorylates sphingosine to form sphingosine 1-phosphate.
XX The polynucleotide was isolated from an HeLa cDNA library by
XX PCR amplification. The invention provides polynucleotides (see
XX CC AAA50508-10) and polypeptides (see AAY96057-59) for the human
XX sphingosine kinase (SK) homologues SKA, SKB and SKC. The
XX polynucleotides may be used as hybridization probes, in the
XX construction of PCR primers for chromosome and gene mapping, in
XX the recombinant production of SKA, SKB and SKC, and in the
XX generation of antisense DNA or RNA. They can be used to detect
XX inflammation or disease associated with abnormal levels of SK
XX expression, or to detect differences in gene sequence between
XX normal and carrier or affected individuals. Host cells expressing
XX SK can be used in drug screening. Human SK specific antibodies,
XX inhibitors, ligands or their analogues can be used as bioactive,
XX agents to treat inflammation or disease including viral, bacterial
XX or fungal infections, allergic responses, mechanical injury
XX associated with trauma, hereditary diseases, lymphoma or carcinoma,
XX and other conditions with activate the genes of kidney, lung,
XX heart, lymphoid or tissues of the nervous system.

SQ Sequence 1447 BP: 256 A; 427 C; 465 G; 299 T; 0 other;

Alignment Scores:

Pred. No.:	6,17e-96	Length:	1447
Score:	106.00	Matches:	143
Percent Similarity:	98.62%	Conservative:	0
Best Local Similarity:	98.62%	Mismatches:	1
Query Match:	27.60%	Indels:	2
DB:	21	Gaps:	0

US-09-784-810A-2 (1-384) x AAA50508 (1-1447)

OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 7 ATGATTCACAGCGGGGGGCCCCCGGGGCGTGTCCCGGCGCTGCTGATGCTG 66
OY 21 LeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
DB 67 CTGAACCCGCGCGGGGCGCAAGGCAAGGCTTTCAGCTTCCGAGATCACCTGACGCC 126
OY 41 LeuLeuAlaGlnAlaGlnLeuLeuSerPheThrLeuMetLeuThrGluArgGlnHisAla 60
DB 127 CTTTGGCTGAGCGCTGAATCTCTTTCACGCTGATGCTCTACTGAGCGGGGCAACGGC 186
OY 61 ArgGluLeuValArgSerGlnGluLeuGlnArgTyrPaspAlaLeuValAlaMetSerGly 80
DB 187 CGGAGAGCTGCGCGGCGGAGAGAGCTGGGCGCCCTGGAGAGCTGTGATGCTGGA 246
OY 81 AspGlyLeuMetHisGlyValValAsnGlyLeuMetGluArgProAspTyrGlnThrAla 100
DB 247 GACGGGCTGATGCTACAGAGGTGTGAACGGGCTCATGAGACGGGCTGACTGGAGACCGC 306
OY 101 IleGlnLysProLeuGlySer-LeuProAlaGlySerGlyAsnAlaLeuAlaSerLe 120

CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 675 BP; 100 A; 234 C; 240 G; 98 T; 3 other;

Alignment Scores:

Pred. No.:	2,91e-65	Length:	675
Score:	75.00	Matches:	75
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	19.53%	Indels:	0
DB:	22	Gaps:	0

US-09-784-810A-2 (1-384) x AAH07214 (1-675)

QY	1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu	20
DB	350 ATGGATCCAGCGGGGGGCGCCCGGGGGCTGCTCCCGCGCCCTGCGGTGCTGCTG	409
QY	21 LeuAsnProArgGlyGlyGlyValLeuGlnLeuPheArgSerHisValGlnPro	40
DB	410 CTGAACCGCGCGGGGCAAGGCGCTTGCAGCTCTTCCGGAGTCACGTGCAGCCC	469
QY	41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla	60
DB	470 CTTTGGCTGAGGCTGAATCTCTTCAAGCTGATGCTCAGCGGCGGAAACACGCG	529
QY	61 ArgGluLeuValArgSerGluLeuGlyArgTrpAspAlaLeu	75
DB	530 CGGAGAGCTGTGCGGTGAGAGAGCTGGCGCTGGACGCTCTT	574

Search completed: July 12, 2003, 19:55:30
 Job time : 301 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: July 12, 2003, 19:15:28 ; Search time 3410 Seconds

(Without alignments)
3277.266 Million cell updates/sec

Title: US-09-784-810A-2
Perfect score: 384
Sequence: 1 MDPAGGPRGVLPKRCVTLV.....CVEPPPSMKPQOMPPEEPL 384

Scoring table:

OLIGO	
Xgapop 60.0 , Ygapext 60.0	
Xgapop 60.0 , Ygapext 60.0	
Xgapop 6.0 , Ygapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2054640 segs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+g2n.model -DEV-xlh
-O/cgn2.1/USPRO.spool/US09784810/runat_11072003_103504_27195/app-query.fasta.1.583
-DB-genemb1 -QFMT-fastap -SUFFIX-olig.rge -MINMATCH-0.1 -LOOPEL-0 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdi -LIST-45
-DOCALLIGN-200 -THR.SCORE-quality -THR.MIN-1 -ALIGN-15 -MODE-LOCAL -OUTFMT-pto
-NORExt-HEARSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US09784810_GCGN.1.1_2496_errunat_11072003_103504_27195 -NCPU-6 -ICPU-3
-NO_MAP -LARGESORT -NEG.SCORES-0 -WAIT -DSPLOCK-100 -LONGLOG
-DEV.TIMEOUT-120 -WARN.TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database :

1: gb_ba:*	
2: gb_hg:*	
3: gb_in:*	
4: gb_om:*	
5: gb_ov:*	
6: gb_pat:*	
7: gb_ph:*	
8: gb_pl:*	
9: gb_pr:*	
10: gb_ro:*	
11: gb_sts:*	
12: gb_sy:*	
13: gb_un:*	
14: gb_vi:*	
15: em_ba:*	
16: em_fun:*	
17: em_hum:*	
18: em_in:*	
19: em_mnu:*	
20: em_om:*	
21: em_or:*	
22: em_ov:*	
23: em_pat:*	
24: em_ph:*	
25: em_pl:*	
26: em_ro:*	
27: em_sts:*	
28: em_un:*	

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_ror:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	384	100.0	1600	6 AX224379	AX224379 Sequence
2	144	37.5	1155	6 AX127642	AX127642 Sequence
3	144	37.5	1173	9 AF200328	AF200328 Homo sapi
4	144	37.5	1693	9 AF266756	AF266756 Homo sapi
5	144	37.5	1719	6 AX127641	AX127641 Sequence
6	144	37.5	1772	9 AK023393	AK023393 Homo sapi
7	144	37.5	1783	9 AF238083	AF238083 Homo sapi
8	144	37.5	1821	9 AK022402	AK022402 Homo sapi
9	144	37.5	1824	9 BC008040	BC008040 Homo sapi
10	141	36.7	1869	9 BC009419	BC009419 Homo sapi
11	141	36.7	2502	9 AK093578	AK093578 Homo sapi
12	125	32.6	2015	9 AB046025	AB046025 Macaca fa
13	115	29.9	1192	9 BC004112	BC004112 Homo sapi
14	115	29.9	183443	2 AC068145	AC068145 Homo sapi
15	115	29.9	209861	2 AC021196	AC021196 Homo sapi
16	98	25.5	1428	9 HSA245504	AJ745504 Homo sapi
17	80	20.8	240	6 AX127662	AX127662 Sequence
18	63	16.4	296	6 AX198352	AX198352 Sequence
19	63	16.4	296	6 AX208818	AX208818 Sequence
20	39	10.2	394	6 AX334274	AX334274 Sequence
21	33	8.6	1146	10 AF415213	AF415213 Mus muscu
22	33	8.6	1559	6 AX287138	AX287138 Sequence
23	33	8.6	1559	10 AF068749	AF068749 Mus muscu
24	33	8.6	1759	6 AX224381	AX224381 Sequence
25	33	8.6	1815	6 AX287137	AX287137 Sequence
26	33	8.6	1815	10 AF068748	AF068748 Mus muscu
27	23	6.0	110000	2 AL606505	AL606505 Mus muscu
28	23	6.0	112037	10 AL645851	AL645851 Mouse DNA
29	22	5.7	1842	10 AB049573	AB049573 Rattus no
30	22	5.7	1895	10 AB049572	AB049572 Rattus no
31	22	5.7	1943	10 AB049571	AB049571 Rattus no
32	22	5.7	2501	10 AB049574	AB049574 Rattus no
33	22	5.7	2648	10 AB049575	AB049575 Rattus no
34	10	2.6	2345	10 BC016109	BC016109 Mus muscu
35	10	2.6	2391	10 RNP1P1NP	X89962 R. norvegicu
36	10	2.6	6618	14 ORSVRNA	X82130 Odonotogloss
37	10	2.6	120537	2 AC096601	AC096601 Rattus no
38	10	2.6	171765	2 AC128377	AC128377 Rattus no
39	10	2.6	196417	2 AC112627	AC112627 Rattus no
40	10	2.6	231664	2 AL845274	AL845274 Mus muscu
41	10	2.6	283318	2 AC102103	AC102103 Mus muscu
42	10	2.6	293272	2 AC120685	AC120685 Rattus no
43	9	2.3	835	3 AD060411	U60411 Anopheles d
44	9	2.3	841	3 AD060410	U60410 Anopheles d
45	9	2.3	1280	8 AY089127	AY089127 Arabidops

RESULT 1

ALIGNMENTS

[illegible]

1 ATGCATCCAGCGGGGGCCCCCGGGGCGTGTCCCGGGCCCTGGCGGTGGTGTG 60
21 LeuasnProArgGlyGlyGlyValAlaLeuGlnPheArgSerHisValGlnPro 40
61 CTGAACCCGCGCGCGGCAAGGCGCTTGCAGCGCTTCCGGAGTCAGGTGAGCC 120
41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
121 CTTTGGCTGAGGCTGAAATCTCTTTCACGCTGATGCTTCACTGACGGCGGAAACCAGCG 180
61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
181 CGGAGCTGTGGCGGTGCGAGAGAGCTGGCGGCTGGAGCGCTGTGGTCATGCTCGGA 240
81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
241 GACGGCTGATGACAGAGGTGTGAACGGCTCATGAGAGCGGCTGAGTGGAGAGCCGCC 300
101 IleGlnLysProLeuCySerLeuProAlaGlySerGlyAspAlaLeuAlaSerLeu 120
301 ATCCAGAGCCCTGTGTAGCTTCCACAGAGCTTGGCAACGCCCTGGACGCTTCTTG 360
121 AsnHisTrpAlaGlyTrpGluGlnValThrAsnGluAspLeuThrAsnCySerLeu 140
361 AACCATATGCTGCTATGACAGAGCTACCAATGAAGACCTCTGACCACTGACGCTA 420
141 LeuLeuCyArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerG 160
421 TTGCTGTGCGC--CGGCG--TGCTGTACACCATGAACCTCTCTCTGTGACACGGCTTGG 478
160 LysLeuArgSer-PheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeu 179
479 GCGTGGCGCT--CTTCTCTGTGCTCAGCCTGGCGCTTCTGCTGATGATGTGACCTA 537
180 GluSerAsp-LysTrpArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuAr 199
538 GAGAGTGA--GAGATATCGCGCTGTGGGAGATATGCGCTTCACTGCGCACCTTCTCGC 596
199 GluAlaAlaLeuArgTrpTrpArgGlyArgLeuAlaThr--LeuProValGlyArgValG 219
597 TCTGGACGCTGTGGCCACCTACCGCGCGGCGGCTGAGCTTCCCTGTAGGAAGAGT--G 654
219 LysPhe-LysThrProAlaSerProValValGlnGlnGlyProValAspAlaHisLeu 238
655 GGTTCACAGACACCTGCTCCCGCTGTGGTCCAGAGGCGCGGTAGATGACACCTT 714
239 ValProLeuGluGluGln--ValProSerHisTrpGln--ValValProAspGluAspPhe 258
715 GTGCCACTGGAGAGACC--AGTGCCCTCTCACTGGAC--AGTGGTGCCGACGAGGACTTGG 772
258 AlLeuValLeuAlaLeuHisSerHisLeuAla--SerGluMetPheAlaAlaPromet 277
773 TGCTAGTCTGTGGACATGCTGCACCTGCACCTGGG--CACTGAGATGTTTGTGCGACCCATG 831
278 GlyArgCysAlaAlaGlyValMetHisLeuPheTrpValAlaGluAlaValSerArgAla 297
832 GCGCGCTGTGCAGCGGTGCATGATCTGTTCTACGTCGCGGCGGAGATGCTGTGCC 891
298 MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTrpGluCysPro 317
892 ATGCTGTGCGCTCTTCTCTGCGCATGAGAAAGGCAAGCATATGAGTATGAAATGCCCC 951
318 TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
952 TACTTGTATATGTGCGGTGTGCTTCCGCTTGGAGCCCAAGAAAGTGGAGTGG 1011
338 PheAlaValAspGlyGluLeuMetValSerGluAlaValGlnGlyValHisProAsn 357
1012 TTTGAGTGGATGGGAATGTGTTGTCAGAGCGCGTCGACAGGCGCAGTGACCCAAAC 1071
358 TyrPheTrpMetValSerGlyCysValGluProProSerTrpLysProGlnMet 377

Db 1072 TACTTGTGATGTACAGCGGTGCTGGAGCCCGCCAGCTGGAAGCCAGAGATG 1131
Qy 378 ProProProGluGluProLeu 384
Db 1132 CCACCGCCGAGAGACCTCTTA 1152
RESULT 3
AF200328
LOCUS AF200328 1173 bp mRNA linear PRI 25-AUG-2000
DEFINITION Homo sapiens sphingosine kinase (SPHK) mRNA, complete cds.
ACCESSION AF200328
VERSION AF200328.1 GI:9909360
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1173)
Pitson,S.M., D'Andrea,R.J., Vandeleur,L., Moretti,P.A., Xia,P.,
Gamble,J.R., Vadas,M.A. and Wattenberg,B.W.
Human sphingosine kinase: purification, molecular cloning and
characterization of the native and recombinant enzymes
Biochem. J. 350 Pt 2, 429-441 (2000)
JOURNAL 2 (bases 1 to 1173)
MEDLINE 20407120
PUBMED 10947957
REFERENCE
AUTHORS Pitson,S.M., D'Andrea,R.J., Vandeleur,L., Moretti,P.A.B., Xia,P.,
Gamble,J.R., Vadas,M.A. and Wattenberg,B.W.
DIRECT SUBMISSION
JOURNAL Submitted (29-OCT-1999) Human Immunology, IMVS, Frome Road,
Adelaide, SA 5000, Australia
FEATURES
source
1..1173
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="vascular endothelial"
1..1173
/gene="SPHK"
19..1173
/gene="SPHK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="AA01980.1"
/db_xref="GI:9909361"
CDS
19..1173
/gene="SPHK"
BASE COUNT 188 a 349 c 390 g 246 t
ORIGIN
Alignment Scores:
Pred. No.: 2,66e-141 length: 1173
Score: 144.00 Matches: 144
Percent Similarity: 100.00%
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 37.50% Indels: 0
DB: 9 Gaps: 0
US-09-784-810A-2 (1-384) x AF200328 (1-1173)
Qy 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
Db 19 ATGCATCCAGCGGGGGCCCCCGGGGCGTGTCCCGCGGCTTGGCGGTGGTGTG 78
Qy 21 LeuasnProArgGlyGlyGlyValAlaLeuGlnPheArgSerHisValGlnPro 40
Db 79 CTGAACCCGCGCGCGGCAAGGCGCTTGCAGCGCTTCCGGAGTCAGGTGAGCC 138
41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60

|||||
Db CTTTGGCTGAGGCTAAATTCCTTCACGCTGATGCTACTAGCGGGAGCAACCGC 198
QY ArgGlutLeuValArgSerGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
Db CCGGAGCTGTGGCGCGGAGGAGCTGGCGCGCTGGACCTCTGGTGGTCATGTCCTGA 258
QY AspGlyLeuMetHisGluValAlaMetGlyLeuMetGlyLysProAspTrpGluThrAla 100
Db GAGCGGCTGATGACAGAGTGGTGAACGGGCTCTATGAGAGCGGCTGACAGAGCGCC 318
QY 11leGlnLysProLeuGlySerLeuProAlaGlySerGlyLysAlaLeuAlaSerLeu 120
Db ATCCAGAGCCCTGCTGTAGCTCCACAGGCTCTGCAACGCGTGGACCTTCCTTG 378
QY 121 AsnHisLysAlaGlyTrpGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeu 140
Db AACCATATGCTGCTATGACGAGGTCAACAAATGAGACCTCTGACCAACTGACGCTA 438
QY 141 LeuLeuGlyArg 144
Db 439 TTGCTGTGCGCC 450
RESULT 4
AF266756 1693 bp mRNA linear PRI 01-JUN-2000
LOCUS AF266756
DEFINITION Homo sapiens sphingosine kinase (SPHK1) mRNA, complete cds.
ACCESSION AF266756
VERSION AF266756.1 GI:8133099
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1693)
Melendez,A.J., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.
Human Sphingosine Kinase, Molecular Cloning, Functional
Characterization and Tissue Distribution
2 (bases 1 to 1693)
Melendez,A.J., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.
Direct Submission
Submitted (11-MAY-2000) Department of Molecular and Cellular
Biology, Jouveinal Parke-Davis, 11-13 rue de la Loge, Fresnes
94265, France
FEATURES
source
1..1693
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q25.2"
1..1693
/gene="SPHK1"
270..1424
/gene="SPHK1"
/codon_start=1
/product="sphingosine kinase"
/protein_id="AAF73470.1"
/translation="MDPAGRGVLPKPCRVILNPRGKGKALQLRSHVOPILAE
AESIFPMTERNHARLAVSEELGEMWDAIYMSDEGIMHVVNGLMERDQETAO
KPLCSIPAGSNAALASLNHYAGYEYTNEDLITNTLLICRLLSPNNLLSTFASG
LRLFSYLSLAWGFIADVDLSEKIRLIGERFTLGTFLALAKRLRGLAVLFVGA
GSKTFASPVVQGVDAHLVLEPVSHWTVDPEDVLALLLSHLSSEFAAP
MGRCAAGVNHLEFYAGVSRAMLRLFLAMEKGRHNEYDCPLTVYPVAAARLEPKDG
KGVFAVDGLMSEAVQGVHNPYFMVNSGVPPPSMKRQMPPEBPL"
1676..1681
/gene="SPHK1"
polyA_signal
BASE COUNT 279 a 529 c 556 g 329 t
ORIGIN
Alignment Scores: 3.57e-141 Length: 1693

Score: 144.00
Percent Similarity: 95.42%
Best Local Similarity: 95.42%
Query Match: 37.50%
DB: 9 Gaps: 0
US-09-784-810A-2 (1-384) x AF266756 (1-1693)
QY 1 MetAspProAlaGlyLysProArgGlyValLeuProArgProCysArgValLeuValLeu 20
Db ATGATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 329
QY 21 LeuAsnProArgGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
Db CTGAACCCGCGCGCGCAAGGCAAGGCTTGGACCTTCCTCGGAGTCACTGACAGGCC 389
QY 41 LeuLeuAlaGluAlaLysSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
Db CTTTGGCTGAGGCTGAATTCCTTCACGCTGATGCTCACTGAGCGGCGGAGAACCGCG 449
QY 61 ArgGluLeuValArgSerGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
Db CCGGAGCTGTGGCGCGGAGGAGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509
QY 81 AspGlyLeuMetHisGluValAlaMetGlyLeuMetGlyLysProAspTrpGluThrAla 100
Db GACGAGCTGATGACAGAGGTGGTGAACGGGCTCTATGAGAGCGGCTGACAGAGCGCC 569
QY 101 11leGlnLysProLeuGlySerLeuProAlaGlySerGlyLysAlaLeuAlaSerLeu 120
Db ATCCAGAGCCCTGCTGTAGCTCCACAGGCTCTGCAACGCGTGGACCTTCCTTG 629
QY 121 AsnHisLysAlaGlyTrpGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeu 140
Db AACCATATGCTGCTATGACGAGGTCAACAAATGAGACCTCTGACCAACTGACGCTA 689
QY 141 LeuLeuGlyArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerG 160
Db TTGCTGTGCGCG-CCGGG-TGCTGTACCAAGAACCGCTGTCTGACAGCGGTTGG 747
QY 160 LysLeuArgSer-PheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeu 179
Db GGCTGCGCCCT-CTTCTCTGCTGCTAGCTGCGGCGGCTTCATTCGATGTGAGCTTA 806
QY 180 GluSerAsp-LysTyrArgArgGlyGluMetArgPheThrLeuGlyTyrPheLeuArg 199
Db GAGAGTGA-GAAGTATCGGCGCTGCGGAGATGCGCTTCACCTGCGGACCTTCTCGG 865
QY 199 GluAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThr-LeuProValGlyArgValG 219
Db TCTGGACGCTCGGACGCTACCGGCGGAGCTGACCTA-CCTCCCTGTAGAGAGAGT-G 923
QY 219 LysPhe-LysThrProAlaSerProValValAlaGlnGlnGlyProValAspAlaHisLeu 238
Db GGTTCAGAGACACTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
QY 239 ValProLeuGlnGlnGln-ValProSerHisTrpGln-ValValProAspGluAspPhe 258
Db GTGCACTGAGAGAGGCG-AGTGCCCTCTCACTGAGC-AGTGCGCCGAGAGAGACTTGG 1041
QY 258 AluLeuValAlaLeuAlaLeuHisSerHisLeuAla-SerGluMetPheAlaAlaProMet 277
Db TGTAGAGCTCGGAGGAGCTGACCTGACCTGCGG-CAGTGAAGVGTTCCTGACCCATG 1100
QY 278 GlyArgCysAlaAlaGlyValMetHisLeuPheThrValAlaArgAlaGlyValSerArgAla 297
Db GCGCGCTGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1160
QY 298 MetLeuLeuArgGluPheLeuAlaMetGluLysGlyArgHisMetGluLysTrpLysPro 317
Db ATGCTGCTGCGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220
QY 318 TyrLeuValLysValProValAlaAlaPheArgLeuGluLysProLysAspGlyLysVal 337

DB 1221 TACTGTGATATGTCGCGTGGCTTCGCTTCGAGCCCAAGAGATGGAGAGGTGTG 1280
QY 338 PhealaValaSPgLygluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsn 357
DB 1281 TTTCGAGTGGATGGGGAATGATGTTAGCCAGAGCCCTGCAGGGCCAGGTCCACCCCAAC 1340
QY 358 TyrPheTpmetValSerGlyCysValGluProProSerTTPlysProGlnMet 377
DB 1341 TACTTCTGATGTCAGCGGTTCGTCGAGCCGCCAGCTGGAGACCCCAAGCATG 1400
QY 378 ProProGluGluProLeu 384
DB 1401 CCACCGCAGAGAGCCCTTA 1421
RESULT 5
AX127641 1719 bp DNA linear PAT 15-MAY-2001
LOCUS Sequence 1 from Patent WO0131029.
ACCESSION AX127641
VERSION AX127641.1 GI:14134307
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Allen,J., Gosink,M., Melendez,A.J. and Takacs,L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 1 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source 1. 1719
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 305 a 529 c 556 g 329 t
ORIGIN
Alignment Scores:
Pred. No.: 3,62e-141 Length: 1719
Score: 144.00 Matches: 375
Percent Similarity: 95.42% Conservative: 0
Best Local Similarity: 95.42% Mismatches: 9
Query Match: 37.50% Indels: 18
DB: 6 Gaps: 0
US-09-784-810a-2 (1-384) x AX127641 (1-1719)
QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 270 ATGATCCAGCGGGGCGGCCCGGGCGTCCCGGGCCCTGCCTGCTGCTGCTG 329
QY 21 LeuAsnProArgGlyGlyGlyGlyValLeuGlnLeuPheArgSerHisValGlnPro 40
DB 330 CTGAACCCGCGCGGGGCAAGGAGGCTTGGACGCTTCCGGAAGTCCAGCTGAGGCC 389
QY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
DB 330 CTTTGGCTGAGGCTGAATCTCTTTCACGCTGATGCTCAGTACGCGGAGACACCGG 449
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTTPAspAlaLeuValAlaMetSerGly 80
DB 450 CGGAGAGTGGTGGCTGGAGAGAGCTGGCGGCTGGAGCGCTGCTGCTGCTGCTGGA 509
QY 81 AspGlyLeuMetHisGluValAlaValAsnGlyLeuMetGluArgProAspTTPGlnThrAla 100
DB 510 GACGGGCTGATGACAGAGTGGTGAACGGCTCATGAGGCGGCTGACTGGGAGACCGCC 569
QY 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyValAsnAlaLeuAlaSerLeu 120
DB 570 ATCCAGAGAGCCCTCTGTAGCTCCACAGAGGCTGTCGACAGCGCTGGACGCTTCTTG 629
QY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeu 140

DB 630 AACCATTTGCTGGCTATGACAGAGTACCAATGAAGACCTCTGAGCAACTGCACGCTA 689
QY 141 LeuLeuCysArgProVal - LeuSerPrometAsnLeuLeuSerLeuHisThrAlaSerG 160
DB 690 TTGCTGTGCTCC - CCGGCG - TGCTGTCACCAAGAACCTCTCTCTGTCACACAGGCTTCGG 747
QY 160 LysLeuArgSer - PheSerValLeuSerLeuAlaTTPGlyPheIleAlaAspValAspLeu 179
DB 748 GGGTGGGCTT - CTCTCTGTGCTGACGCTGGGCTGGGGCTTTCATTTGATGTGAGCTTA 806
QY 180 GluSerAsp - LysTyrArgArgLeuGlyGlyLeuMetArgPheThrLeuGlyThrPheLeuArg 199
DB 807 GAGAGTGA - GAAAGTATGGCGCTGCGGGAGATGGCGTTCATCTGGGACCTTCCTCGC 865
QY 199 GLeuAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThr - LeuProValGlyArgValG 219
DB 866 TCTGCGACCTCTGCGCACCTACCGCGCGGCGAGCTGCCTTA - CCTCCCTGTAGGAAGAGT - G 923
QY 219 LysPhe - LysThrProAlaSerProValValAlaGlnGlnGlyProValAspAlaHisLeu 238
DB 924 GCTTCCAGACACCTGCTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983
QY 239 ValProLeuGluGluGln - ValProSerHisTTPGln - ValValProAspLysAspPhe 258
DB 984 GTGCCACTGGAGAGACC - AGTGCCCTCTCAGCTGAGC - AGTGAGTGGCCGACGAGACCTTGG 1041
QY 258 alleuValLeuAlaLeuLeuHisSerHisLeuAla - SerGluMetPheAlaAlaProMet 277
DB 1042 TGCTGATGCTGTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
QY 278 GlyArgCysAlaAlaGlyValAlaMetHisLeuPheThrValArgAlaGlyValSerArgAla 297
DB 1101 GCCCGCTGTCAGCTGGCGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160
QY 298 MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysArgPro 317
DB 1161 ATGCTGTCGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1220
QY 318 TyrLeuValTyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
DB 1221 TACTGTGATATGTCGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1280
QY 338 PhealaValaSPgLygluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsn 357
DB 1281 TTTCGAGTGGATGGGGAATGATGTTAGCCAGAGCCCTGCAGGGCCAGGTCCACCCCAAC 1340
QY 358 TyrPheTpmetValSerGlyCysValGluProProSerTTPlysProGlnMet 377
DB 1341 TACTTCTGATGTCAGCGGTTCGTCGAGCCGCCAGCTGGAGACCCCAAGCATG 1400
QY 378 ProProGluGluProLeu 384
DB 1401 CCACCGCAGAGAGCCCTTA 1421
RESULT 6
AK023393 1772 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ13331 f1s, clone OVARC1001809, moderately
DEFINITION Similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
ACCESSION AK023393
VERSION AK023393.1 GI:10435311
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens ovary, tumor tissue cDNA to mRNA, clone_11b-OVARC1
clone:OVARC1001809.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosofiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,

Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahashi, K., Masuho, Y., Niinomiya, K., and Iwayanagi, T. NEDO human cDNA sequencing project
JOURNAL
TITLE
REFERENCE
AUTHORS
JOURNAL
2 (bases 1 to 1772)
Isogai, T. and Otsuki, T.
Submitted (23-NOV-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@helix.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES
Source
Location/Qualifiers
1..1772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC1001809"
/tissue="ovary, tumor tissue"
/clone_lib="OVARC1"
/note="cloning vector: PME18SFL3"
350..1504
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14558.1"
/db_xref="GI:10435312"
/translation="MDPAGRGVLPAPPCRVILNPRGKGKALQLPFRSHVPLLEAEISPLMTERRNHRARLVRSEELGADLVVMSGDLHMFVNGVLMWRDMEFLAIGKPLSPASSGALAAASHNRYAGEVQNTEDLLQNTLLCRLLSPNNLSTHRSGLRFSVLSIANGFADVCLSEKRRRLGEMRPGLRLLALRYRRLATVYGVY GSKTPASPVYVQCGVDHLYPLLEPPVSHWTVPDEYVLLALHSHLSSEMPAR MGRCAAGVNHLETVRAGVSRAMLRLFLAMEKGHHEIRPFLVTVYVAFLERKDS KGVFAVDGMLNVSFAVGVHNFVMSGCVPEPPSMKFPQMPPEEPFL"
BASE COUNT 285 a 566 c 585 g 336 t
ORIGIN
Alignment Scores:
Pred. No.: 3,71e-141 Length: 1772
Score: 144.00 Matches: 375
Percent Similarity: 95.42% Conservative: 0
Best Local Similarity: 95.42% Mismatches: 9
Query Match: 37.50% Indels: 18
DB: Gaps: 0
US-09-784-810a-2 (1-384) x AK023393 (1-1772)
QY 1 MetaspProlaaglylProargglyValleuProlarProcySarvValleuValleu 20
DB 350 ATGATCCACGCGGCGCCCGGCGCGTGTCCCGCGCGCGCGCGTGTGTGTG 409
QY 21 LeuansProlarglylLysglyLysAlaLeuGlnLeuPheargSerHisValGlnPro 40
DB 410 CTGAACCCGCGCGCGCAAGGCAAGGCTTTCAGCTTCGCGAGTCACTGCGAGCCC 469
QY 41 LeuLeuAlaGlnAlaGlnLeuSerPheThrLeuMetLeuThrGlnUArgAArgAsnHisAla 60
DB 470 CTTTGGCTGAGGCTGAATCTCTTCACGCGTCACTGAGCGCGCGCAACCAACCG 529
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValAlaMetSerGly 80
DB 530 CGGAGCTGTGCGCGCGAGGAGCTGGGCGCTGGGAGCGCTGTGGTCACTGTGGA 589
QY 81 AspGlyLeuMetHisGlnValAlaAsnGlyLeuMetGlnUArgProAspTrpGlnHisAla 100
DB 590 GACGGCTGATGCACGAGGTGTGAACGGGCTCATGAGCGGCTGACTGGAGACCGCC 649
QY 101 IleglnLysProlLeuGlySerLeuProlaaglylSerGlyAsnAlaLeuAlaSerLeu 120

DB 650 ATCCAGAGCCCTGTGTAGCCCTCCACAGAGCTCTGGCAAGCGCGTGGACCTTCTTG 709
QY 121 AsnHisTyrAlaGlyTyrGlnGlnAlaValThrAsnGlnUAspLeuThrAsnGlyThrLeu 140
DB 710 AACCATTAATGCTGGTATAGAGAGGTCCCAATGAAAGACTCTGACCACTGACAGCTA 769
QY 141 LeuLeuGlyArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerG 160
DB 770 TTGCTGTGGCCG--CCGGC--TGCTGTACCACTGAACCTGCTGTCTTGCACAGCGTTGG 827
QY 160 LysLeuArgSer--PheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeu 179
DB 828 GGGCGCCCGC--CTTCTCTGTGTCTAGCCCTGGCGGGGCTTCACTGTGATGTGACCTTA 886
QY 180 GluSerAsp--LysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuAr 199
DB 887 GAGAGTGA--GAAATATGCGCGCTGTGGAGATGCGCTTCACTGTGCGACCTTCTTCCG 945
QY 199 GLeuAlaAlaLeuAlaArgThrTyrArgGlyArgLeuAlaThr--LeuProValGlyArgValG 219
DB 946 TCTGGCAGCCCTCGGCACTACCGCGCGCACTGGCTA--CTTCCCTGTAGAGAGAT--G 1003
QY 219 LysPhe--LysThrProAlaSerProValAlaGlnGlnGlnProValAspAlaHisLeu 238
DB 1004 GGTTCACAGACACCTGCTCCCGCTTGTGTGTCCAGAGGCGCCGTTAGATCAACACTT 1063
QY 239 ValProlLeuGluGlnGln--ValProSerHisTrpGln--ValValProAspGluAspPheY 258
DB 1064 GTGCCACTGGAGAGAGCC--AGTCCCTCTCACTGTGAC--AGTGTGCCGACAGAGACTTGG 1121
QY 258 allLeuValAlaLeuAlaLeuLeuHisSerHisLeuAla--SerGluMetPheAlaAlaProMet 277
DB 1122 TGCTAGTCCCTGGAGACTGTGCACCTGCACCTGGG--CAGTAAAGATGTTCTCTGACCCATG 1180
QY 278 GlyArgGlyAlaAlaGlyAlaValMetHisLeuPheTyrAlaArgAlaGlyAlaSerArgAla 297
DB 1181 GCGCCGCTGCACCTGCTGCTCATGCTGTCTTCACTGCGGGCGGAGATGCTCTGTGCC 1240
QY 298 MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluLysGluLysPro 317
DB 1241 ATGCTGTGCGCTCTCTCTGCGCATGGAAGGCGAGGATGTGATGAATGAATGAAATGCC 1300
QY 318 TyrLeuValTyrValProValAlaAlaPheArgLeuGluProLysAspGlyLysGlyVal 337
DB 1301 TACTTGTATACCTGCGCGTGTGCTGCTGCGCTTGGAGGCCCAAGATGGAAGAGTGTG 1360
QY 338 PheAlaValAspGlyGluLeuMetValSerGlnAlaValGlnGlyGlnValHisProAsn 357
DB 1361 TTTCAGTGGATGGGGAATTGATGTTAGCGAGCGCGTGCAGGCGCAGGTGCACCCAAAC 1420
QY 358 TyrPheThrMetValSerGlyGlyValGlnUArgProProSerTrpLysProGlnGlnMet 377
DB 1421 TACTTCTGGATGCTCAGCGGTTGCTGTGAAGCCCGCCGCAAGCTGGAAGCCCAAGCATG 1480
QY 378 ProProProGluGluProLeu 384
DB 1481 CCACCGCCAGAGAGGCTTGA 1501
RESULT 7
AF238083 1783 bp mRNA linear PRI 01-JUN-2000
LOCUS Homo sapiens sphingosine kinase-1 mRNA, complete cds.
DEFINITION AF238083
ACCESSION AF238083
VERSION AF238083.1 GI:8132867
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1783)
Nava, V.E., Iacana, E., Poulton, S., Liu, H., Sugiyama, M., Kono, K., Milstien, S., Kohama, T., and Spiegel, S.
Functional characterization of human sphingosine kinase-1
TITLE

JOURNAL FEBS Lett. 473 (1), 81-84 (2000)
 MEDLINE 20263733
 PUBMED 10802064
 REFERENCE 2 (bases 1 to 1783)
 AUTHORS Naya,V.E., Lacana,E., Poulton,S., Liu,H., Sugitara,M., Kono,K.,
 Milstien,S., Kohama,T. and Spiegel,S.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-2000) Biochemistry, Georgetown University, 3900
 Reservoir RD NW BSB Rm 357, Washington, DC 20007, USA

FEATURES
 source
 location/Qualifiers
 1. 1783
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q25.2"
 337. 1491
 /codon_start=1
 /product="sphingosine kinase-1"
 /protein_id="AAF73423.1"
 /db_xref="GI:8132868"
 /translation="MDPAGRGVLPKRCVYLILNPRGKALQIFRSHVOPILLAE
 AISTTLTERRNHARLVSEELGRMDALVYMGDGLMEHYNGLMERDMEATIO
 KPLCSLPAGSGNALAASLNHYAGYEQVYNEDLLNCTLLLRLLSPNNLSLHTASG
 LRFSVLSLWGFADVDLESEKRLDEEMFTLTGTLALRLYRGLAYLPVGRV
 GSKTPASPVVVOOGVDNALVLEPVPSTHMTVDEDFVLIALHSHLSSEMPAP
 MGRCAGYMHLFYRAGVSRAMLRLFLAMEGRMECEPVLVYVAVARLEPKDG
 KGVFAVDDELAVSEAVOGVHPNFTMYSGCEVPEPPSKPQMPPEEPL"

CDS

BASE COUNT 314 a 546 c 591 g 332 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,79e-141 Length: 1783
 Score: 144.00 Matches: 144
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 37.50% Indels: 0
 DB: 9 Gaps: 0

US-09-784-810A-2 (1-384) x AF238083 (1-1783)

QY 1 Metaspproalaglyglyproarglyvalleuproargprocyargvalleuvalleu 20
 DB 337 ATGATTCACAGCGGGGCGCCCGGGCGTCTCCCGCGCCCTGCGGTGCTGCTG 396
 QY 21 LeuAanproargglyglylysalyalaenuglnleupheargserhisvalglnpro 40
 DB 397 CTGAACCCGCGGGGCGGCAAGGCGCTTGCAGCTCTTCCGAGTCACGTCAGCC 456
 QY 41 LeuAanlaglualagluileserpherthleumetleuthrghlurgarganhsala 60
 DB 457 CTTTGGCTGAGGCTGAATCTCTTCAAGCTGATGCTGACGACGCGGCGAGCACGCG 516
 QY 61 Arggluleuvalargserglucluleuglyargtrpaspalaaleuvalalemsersely 80
 DB 517 CGGGAGCTGCTCGCTCGAGGAGCTGGCGGCGCTGCTGCTGCTGCTGCTGCTG 576
 QY 81 AspglyleuAmetHisgluValaValasnglyleuMetgluargproasprtrgluThrAla 100
 DB 577 GACGGGCTGATGACAGAGGTGGTGAAGGCTCATGAGAGGCGCTGAGTGGAGACCGCC 636
 QY 101 IleaglinsproleucysserleuproalaglyserglyasalaaleuAlaAlaSerleu 120
 DB 637 ATCCGAAGCCCTGTGAGCTCTCCACAGGCTTGGCAACGCGCTGCGACCTTCCCTTG 696
 QY 121 AsnhslyrAlaglyTyrGlulnValThrasngluaspLeuThrAsnGysThrleu 140
 DB 697 AACCCATTATGCTGCTATGAGCAGCTCACCAATGAAGACCTCTGACCAACTGACGCTA 756
 QY 141 LeuAanCysArg 144
 DB 757 TTGCTGTGCTGCTG 768
 RESULT 8

AK022402 1821 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens CDNA FLJ12340.f1s, clone MAMMA1002268, moderately
 DEFINITION similar to Mus musculus sphingosine kinase (SPK1a) mRNA.
 ACCESSION AK022402
 VERSION AK022402.1 GI:10433790
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens Mammary gland cDNA to mRNA, clone_11b:MAMMA1
 clone:MAMMA1002268.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
 Magatsuma,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
 Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
 Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
 Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakematsu,A.,
 Nakamura,Y., Nagahara,K., Masuo,Y., Nishimura,K. and Iwayanagi,T.
 NEO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 1821)
 Isogai,T. and Otsuki,T.
 Direct Submission
 Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT

NEO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction, 5'- and 3'-end one pass sequencing and clone selection:
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.
 Location/Qualifiers
 1. 1821
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MAMMA1002268"
 /tissue_type="Mammary gland"
 /clone_11b="MAMMA1"
 /note="Cloning vector: pME18SFL3"
 393. 1547
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB14028.1"
 /db_xref="GI:10433791"
 /translation="MDPAGRGVLPKRCVYLILNPRGKALQIFRSHVOPILLAE
 AISTTLTERRNHARLVSEELGRMDALVYMGDGLMEHYNGLMERDMEATIO
 KPLCSLPAGSGNALAASLNHYAGYEQVYNEDLLNCTLLLRLLSPNNLSLHTASG
 LRFSVLSLWGFADVDLESEKRLDEEMFTLTGTLALRLYRGLAYLPVGRV
 GSKTPASPVVVOOGVDNALVLEPVPSTHMTVDEDFVLIALHSHLSSEMPAP
 MGRCAGYMHLFYRAGVSRAMLRLFLAMEGRMECEPVLVYVAVARLEPKDG
 KGVFAVDDELAVSEAVOGVHPNFTMYSGCEVPEPPSKPQMPPEEPL"

FEATURES

source

CDS

BASE COUNT 301 a 569 c 614 g 337 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,79e-141 Length: 1821
 Score: 144.00 Matches: 375
 Percent Similarity: 95.42% Conservative: 0
 Best Local Similarity: 95.42% Mismatches: 9
 Query Match: 37.50% Indels: 18
 DB: 9 Gaps: 0

US-09-784-810A-2 (1-384) x AK022402 (1-1821)

QY 1 Metaspproalaglyglyproarglyvalleuproargprocyargvalleuvalleu 20
 DB 393 ATGATTCACAGCGGGGCGCCCGGGCGTCTCCCGCGCCCTGCGGTGCTGCTG 452
 QY 21 LeuAanproargglyglylysalyalaenuglnleupheargserhisvalglnpro 40

```

Db      453  CTCGACCCGCGCGGCGGCAAGGCGCTTCGACGCTCCGAGACAGCTGCACGCC 512
      41  LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
      513  CTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTACTGAGCGGAGACACCGG 572
      61  ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValaMetSerGly 80
      573  CGGAGCTGGTGGGTCGAGAGAGAGCTGGGCGCTGGAGCGCTGGTGTGATGCTGGA 632
      81  AspGlyLeuMetHisGluValaValaAsnGlyLeuMetGluArgProAspIleuThrAla 100
      633  GAGCGGGTGATGCGAGAGTGTGGAACGGGCTCATGAGCGGCGCTGACTGGAGAGACGCC 692
      101  TleGlnLysProLeuLysSerLeuProAlaGlySerGlyAsnAlaLeuAlaLaserLeu 120
      693  ATCCAGAAAGCCCTGTGTAGCTCCAGCAGGCTCGGCAACGCGCTGGCAGCTTCCTTG 752
      121  AsnHisTyrAlaGlyTyrGluGlnValaThrAsnGluAspLeuThrAsnGlyThrLeu 140
      753  AACCATATGCTGGCTATGACAGAGTACCAATGAAGACCTCCAGACCAACTGCACGCTA 812
      141  LeuLeuGlyArgProVala--LeuSerProMetAsnLeuLeuSerLeuHisTyrAlaSerG 160
      813  TTGCTGTGCGC-CGGGC-TGCTGTACCCCATGAACCTGTGTCTCTGCACACGCGCTTGG 870
      160  LysLeuArgSer-PheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValaAspLeu 179
      871  GGGTGGCGCT-CTTCTCTGTCTGACGCTGGCGCTGGGCTTCATCTGCTGATGTGGACCTA 929
      180  GluSerAsp-LysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArg 199
      930  GAGAGTGA-GAAGTATCGGCGCTGGGGAGATGGCTTCACTCGGGCAGCTTCGCG 988
      199  GLeuAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThr-LeuProValaGlyArgVala 219
      989  TCTGGGAGCCCTCGGACCTACCGCGCGGCTGCGCTCA-CCTCCCTGTAGAGAGAGT-G 1046
      219  LysPhe-LysThrProAlaSerProValaValaGlnGlnGlyProValaAspAlaHisLeu 238
      1047  GGTTCAGAGACACTGGCTCCCGCTGGTGTGTCACAGAGCGCGGTAGATGACACCTT 1106
      239  ValProLeuGluGluGln-ValProSerHisTrpGln-ValValProAspGluAspPheV 258
      1107  GTGCACACTGGAGGAGGC-AGTGGCCTCTCTACTGGAC-AGTGGTGGCGGAGGACTTGG 1164
      258  alleuValLeuAlaLeuLeuHisSerHisLeuAla-SerGluMetPheAlaLapProMet 277
      1165  TGTATGCTCTGGCACTGCTGCACCTGCACCTGGG-CAGTGAAGATGTTGCTGCACCCATG 1223
      278  GlyArgGlyAlaAlaGlyValaMetHisLeuPheThrValaArgAlaGlyValaSerArgAla 297
      1224  GGCCGCTGTGCGAGCTGGCTCATCTGTTCTACTGCGGGCGGAGGTCTCTGTCGCC 1283
      298  MetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGlyGlyPro 317
      1284  ATGCTGCTCGGCTCTTCTGCGCATGGAAGAGCGCATATGAGTGTGATGATGCCCC 1343
      318  TyrLeuValTyrValProValaValaLapPheArgLeuGluProLysAspGlyLysGlyVala 337
      1344  TACTGTGATATGTGCGCTGTGCTTCCTTCGCTGGAGCCCAAGATGGGAAAGTGTG 1403
      338  PheAlaValaAspGlyGluLeuMetValSerGluAlaValaGlnGlyGlnValaHisProAsn 357
      1404  TTTCGAGTGGATGGGAATTGATGTTAGTCGAGGCGCGAGGCGCAGGTGCACCCAAAC 1463
      358  TyrPheTrpMetValSerGlyValaGluProProSerTrpLysProGlnGlnMet 377
      1464  TACTTCTGATGTGCTAGCGGTGTGCTGGAGCCCGCCAGCTGGAAGACCCAGAGAGATG 1523
      378  ProProGluGluProLeu 384

```

Db 1524 CCACCGCCAGAAAGCCCTTA 1544

RESULT 9
LOCUS BC008040 1824 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, clone MGC:15041 IMAGE:3831657, mRNA, complete cds.
ACCESSION BC008040
VERSION BC008040.1 GI:14165485
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 1824)
Strausberg, R.
Direct Submission
Submitted (21-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC/DCMP/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amandansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kerteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 24 Row: n Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10433790.
Location/Qualifiers

FEATURES

1..1824

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:15041 IMAGE:3831657"

/tissue_type="Skin, melanotic melanoma."

/clone_1lb="NIH-MGC_20"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

339..1493

/codon_start=1

/product="Unknown (protein for MGC:15041)"

/protein_id="AAH08040.1"

/db_xref="GI:14165486"

/translation="MDPAGRGVCULPRRCRYULILNPRGKRALQLFRSHVPLAE
AEISFTMLTERNNHARELVRSEELGRWDALVYMGDDLMHEVYNGLMERPDMEATIG
KPLCSLPAGSGNALASLNHYAGYDQVNEDELITCTLLLRLLSPNRLISHTAG
LRLFSYLSLAMEIADVDLESEKRYRIGLMTPLDTFLLALRTYRRLAYLRPGV
GSKTRPASPVYVOOGPYVDANLVRLEEVPSHMTVVVDEDFVLVLLSHLSSEFAPR
MGRADAVMLLRYRAGVSRAMLRLFLAMEGRMEYECPLVLYVVRARLRPKOG
KGVYAVDELMSVSEAYVOGYNRYFMVMSGCVPRPSMKRPOOMPRPEBPL"

CDS

BASE COUNT

354 a 551 c 592 g 327 t

ORIGIN

Alignment Scores: 3.8e-141 Length: 1824
Pred. No.: 144.00 Matches: 144
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 37.50% Gaps: 0
Query Match: 9

US-09-784-810A-2 (1-384) x BC008040 (1-1824)

1 MetAspProAlaGlyGlyProAlaGlyValaLeuProArgProCysArgValaLeu 20
|||||

Db 339 ATGATCCAGCGGGGCCCCGGGGGCTGCTCCCGGCCCTGCTGCTGCTGCTG 398
Qy 21 Leuansproargglylyglylyglylyalaleuclnleupheargserhshvalgnpro 40
Db 399 CTGAACCCGGCGGGGCAAGGGGCAAGCCCTTGAGCTCTCCGAGTCACGTGACCCC 458
Qy 41 Leuclnleuaglualeuclnleuclnleuclnleuclnleuclnleuclnleucln 60
Db 459 CTTTGGCTGAGGCTGAATCTCTTACCGCTGATGCTGCTGCTGCTGCTGCTG 518
Qy 61 Arggluleuvalargsergluleuclnleuclnleuclnleuclnleuclnleucln 80
Db 519 CGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 578
Qy 81 Aspdlleuclnleuclnleuclnleuclnleuclnleuclnleuclnleuclnleucln 100
Db 579 GACGGCTGATGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
Qy 101 Illeuclnleuclnleuclnleuclnleuclnleuclnleuclnleuclnleuclnleucln 120
Db 639 ATCCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
Qy 121 AsnHsTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeu 140
Db 699 AACCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 758
Qy 141 Leuclnleuclnleuclnleuclnleuclnleuclnleuclnleuclnleuclnleucln 144
Db 759 TTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 770

RESULT 10
BC009419 1869 bp mRNA linear PRI 12-JUL-2001
LOCUS BC009419
DEFINITION Homo sapiens, clone MGC:15040 IMAGE:3506924, mRNA, complete cds.
ACCESSION BC009419.1 GI:14495624
VERSION MGC.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1869)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DPF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Gulin, Letitia Hsiao, Martin Kirzylinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 23 Row: e Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10433790.
Location/Qualifiers
1..1869
SOURCE

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:15040 IMAGE:3506924"
/class="type="Kidney, renal cell adenocarcinoma"
/clone_id="N1H_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
381..1577
/codon_start=1
/product="Unknown (protein for MGC:15040)"
/protein_id="AAH09419.1"
/db_xref="GI:14495625"
/translation="MDPVGCGRGLFEGFVSAGGRGLPRCPRLVLLNPRGKGA
IOLFRSHVOPULAEATISPLMTERNHARRELYRSRELGRLMALVYMGSDGLMHEV
NGLMERPDWETALOKPLCSIPAGSGNALSLNRYAYEOTYJEDLLTNTLLRL
LSPNNLUSLHTASGLRLESVLSLWAGTIAVDLESEYTRRLEGRFLTGFLALALR
TYRGLAYLLEPVGRGSKTPASPVVVOGPDADHLVPLEEPPSHMTVPDEDFVLA
LLSHSLSEMFAPRGMRCAGVYMLFVRAGVSRAMLLRLFLAMEKGRHMEYDCPYLY
YVPVAPRLERPKDKGVFAVDGELMVAEAVGVHPYFMVVSQVPEPPSMKPOQMP
PPEEPL"

BASE COUNT 322 a 570 c 623 g 354 t
ORIGIN

Alignment Scores:
Pred. No.: 5.54e-138 Length: 1869
Score: 141.00 Matches: 372
Percent Similarity: 95.38% Mismatches: 0
Best Local Similarity: 95.38% Gaps: 18
Query Match: 36.72% Indels: 18
Gaps: 0

US-09-784-810a-2 (1-384) x BC009419 (1-1869)

Qy 4 AlaGlyLeuProArgGlyValLeuProArgProCysArgValLeuValLeuAsnPro 23
Db 432 GCGGGCGGCCCCCGGGGCTGCTCCCGGCCCTGCGCGCTGCTGCTGCTGCTGCTG 491
Qy 24 ArgGlyLeuGlySerPheThrLeuMetLeuThrGluArgAsnHisAlaArgGlyLeu 43
Db 492 CGCGGGCGGCAAGGCAAGGCTTCCAGCTCTCCGAGTCCAGTCCAGTCCAGTCCAGT 551
Qy 44 GluAlaGluLeuSerPheThrLeuMetLeuThrGluArgAsnHisAlaArgGlyLeu 63
Db 552 GAGGCTGAATCTCTTACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611
Qy 64 ValArgSerGluGluLeuGlyArgTyrPaspAlaLeuValAlaMetSerGlyAspGlyLeu 83
Db 612 GTGGGTCGCGAGGAGCTGGGGCGGCTGAGACCTGTGTGTCATGTCTGGAGAGGCGTG 671
Qy 84 MethIsgValValAsnGlyLeuMetGluArgProAspTyrPgluThrAlaIleGlnLys 103
Db 672 ATGCACGAGTGTGGTGAACGGGCTCATGAGAGGGGCTGACTGGAGACCGCATCCAGAG 731
Qy 104 ProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaLeuSerLeuAsnHisTyr 123
Db 732 CCCCTGTGTACCTCCAGCAGGCTGTGGCAACCGGCTGGAGACTTCTTGAACCATATAT 791
Qy 124 AlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeuLeuLys 143
Db 792 GCTGCTATGAGCGAGGTCACCAATGAAGACTCTCTGACCAACTCAGCATTTCTGTGC 851
Qy 144 ArgProVal--LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArg 163
Db 852 CG-CCGGC-TGCTGTACACCATGTAACCTGCTGCTGTGCACACGGCTTGGGGCTGGCC 909
Qy 163 er-PheSerValLeuSerLeuAlaTyrPylPheIleAlaAspValAspLeuGluSerAsp 182
Db 910 T-CTTCTCTGTGTCACAGCTGCTGGGGCTTCAFTGTGTATGGAGCTAGAGAGGA- 967
Qy 183 -LysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeuAla 202
Db 968 GAAGTATCGGCTGCTGGGGAGATGCGCTCATCTGTGGGCACTTCTGCTGCGAGC 1027

QY	242	GLUIGLUGln-ValProSerHisTrpGln-ValValProArgGluAspPheValLeuValL	261
Db	1801	GAGAGACC-AGTGGCCCTCTCACTGGAC-AGTGGTGGCCGACGAGACTTTTGTCTAGTC	1858
QY	261	euaIaleuLeuHisSerHisLeuAla-SerGluMetPheAlaIaIaProMetGlyArgCys	280
Db	1859	TGGCAGCTGCTGCATGCCACCTGGG-CAGTGAAGTGTTCGTGCACCATGGGCGGGTGT	1917
QY	281	AlaIaIaGlyValMetHisIleuPheTyrValArgAlaGlyValSerAlaGluAlaMetLeuLeu	300
Db	1918	GCACCTGGCGGCAAGCATCTGTCTAGCTGGGGGGGAGGGAGTCTCGGCAATCTCTGTG	1977
QY	301	ArgIleuPheLeuAlaMetGlyLysGlyArgHisMetGluTyrGlyCysProTyrLeuVal	320
Db	1978	CGCCTTCTCCGTGGCCATGGAGAAAGGACGAGCATGTGAGTGAATGGACCTTCTGTGTA	2037
QY	321	TyrValProValValAlaIaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal	340
Db	2038	TACCTGGCCGGTGGTGGCTTCCGCTTGGAGCCCAAGAGTGGGAAAGGTGTCTTGGCACTG	2097
QY	341	AspGlyLysLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrp	360
Db	2098	GATGGGGAATTGATGGTTAGCGAGAGCCGTGCAGGGCCAGGTGCACCCAAATTAATTCTGG	2157
QY	361	MetValSerGlyCysValGluProProSerTrpLysProGlnGlnMetProProPro	380
Db	2158	ATGTGTACGGGTTCGTTGAGACCCCCGGCCAGCTGGAAGCCCCAGACATGCACCGGCA	2217
QY	381	GluGluProLeu 384	
Db	2218	GAAAGAGCCCTTA 2229	

LOCUS	2015 bp	mRNA	linear	PRI 18-JUL-2007
DEFINITION	Macaca fascicularis brain cDNA, clone:QcCE-16602.			
ACCESSION	AB046025			
VERSION	AB046025.1 GI:9280105			
KEYWORDS	11s (full insert sequence).			
SOURCE	Macaca fascicularis adult cDNA to mRNA, clone_11b:macaque brain cDNA library QcCE clone:QcCE-16602.			
ORGANISM	Macaca fascicularis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.			
REFERENCE	1 (sites)			
AUTHORS	Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M., Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.			
TITLE	Isolation of full-length cDNA clones from macaque brain cDNA libraries			
JOURNAL	unpublished			
REFERENCE	2 (bases 1 to 2015)			
AUTHORS	Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.			
TITLE	Direct Submmission			
JOURNAL	Submitted (14-JUL-2000) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan			
COMMENT	(E-mail:hashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/, Tel.:81-3-52851111(ex.2120), Fax:81-3-52851181) URL: http://www.nih.go.jp/yoken/genebank/			

Vector: pME18s-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CACCCTGGG)
R. Site2: DraIII (CACCATGCG)
Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGTCGCCCCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18s-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries

were constructed by Sugano et al. (University of Tokyo, Institute of Medical Science). Custom primer used for sequencing (5' and primer (CTTCTCCTCTAAAGCTGG); 3' and primer (CGACCTGCAGCTCAGCACA)).

Location/Qualifiers

1..2015

```

/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="QcCE-16602"
/clone_1lb="macaque brain cDNA library QcCE"
/dev_stage="adult"
628..1617

```

```

/product="unnamed protein product"
/protein_id="BAB01607.1"
/db_xref="GI:9280106"
/translation="MDPAGGREGVLPSCRYVLLNPRGGKALQLFSHVODPLAE
AEIFTLMTERRNHARELVSEBELGRMDALVYVSGGLMEHVEVNGMLSEPDWTAIO
KPLCSLPAGSGNALAALSNHYAGVEYQNTEDLLTNCRIICRLLSPMNLISHTAGS
LRLEFVSLAMGFIADYDVLSEKRYRGEARFTLGTLRLAALNTYRGRLLATLPYGR
GSKTPSPVYVVOGQVDALHVLPLEEPVSHTWDDFVYVLLALSHLSESEFAAP
MGRCAACVMHLEFYRAGVSRAMLRLRLAMEKGRHMEYECPLVYISGWGIDG"

```

Alignment Scores:	
Pred. No.:	3.98e-121
Score:	125.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	32.55%
DB:	9
	Gaps:
	0
	Length:
	2015
	Matches:
	125
	Conservative:
	0
	Mismatches:
	0
	Indels:
	0
	Gaps:
	0

Oy	15	CysArgValLeuValLeuLeuAsnProArgIlgLyLysGlyLysAlaLeuGlnLeuPhe	34
Dd	670	TGCGCGTGCTGGTGGCTGTCTAAACCCGGGTGGCGGCAAGGCCCTCAGCTTTC	729
Oy	35	ArgSerHisValGlnProLeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThr	54
Dd	730	CGAGATCAGCTGCAGACCCCTTTTGTGGCTGAGGCTGAATTCCTTTCAGAGCTGATCTCAGCT	789
Oy	55	GluArgArgAsnHisAlaIleArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAla	74
Dd	790	GAGCGGGGAAACCAATGGCGGGGAGACTGTGTGGTGGCGGAGAGCTGGGCGCTGGGAGAGCT	849
Oy	75	LeuValValMetSerArgLysAspGlyLeuMetHisGluValValAsnGluLeuMetGlnArg	94
Dd	850	CTGGTGGTCATGTGTGGAGACGAACTGATGCACAGAGTGTGAACCGGGCTATGAGACGG	909
Oy	95	ProAspTrpGlnThrAlaIleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsn	114
Dd	910	CTTCACTGGGAGACCGGCATCCAAAGGCCCTGTGTACCTCCCAAGAGGCTTGGGCAC	969
Oy	115	AlaLeuAlaAlaSerLeuAsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeu	134
Dd	970	GCAGTGGGAGCTTCTTGAACCAATTATGCTGGTATGAGCAGGTGCACCAATGAAGACCTC	1029
Oy	135	LeuThrAsnCysThr	139
Dd	1030	CTTGACCAACTGCACG	1044

RESULT 13	
BC004112	
LOCUS	1192 bp mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, similar to sphingosine kinase 1, clone IMAGE:3832587,
ACCESSION	BC004112
VERSION	BC004112
KEYWORDS	BC004112.1 GI:13278662
SOURCE	.
ORGANISM	Homo sapiens.
	Homo sapiens

REFERENCE
TITLE
AUTHORS
JOURNAL

1 (bases 1 to 1192)

Strausberg, R.
Direct Submission
Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC/DCID/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahy, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 14 Row: f Column: 13.

FEATURES
Source

1. 1192
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3832587"
/tissue_type="Skin, melanotic melanoma."
/clone_id="NIH-MGC-20"
/lab_host="DH10B-R"
/note="vector: pOTB7"
1. 873
/codon_start=1
/product="Similar to sphingosine kinase 1"
/protein_id="AAH04112.1"
/db_xref="GI:13278663"
/translation="GTRETAIQKPLCSIPAGSNAALASLNHYAGYEQVNTMEDLTNC
TLICRLSLPMNLISHTASGIRLFSVLSIANGFIADVDSESKYRLGEMRTLT
FLRLALNTYKGRLAYLDYGVGSKTPSPYVVOGDALVPLEPVPBMTVVVD
EDFLVIALHSHGSEMFAPMGKAGVNHLYVRAGVSRALRLFLAMEGRKH
EECEYLVVVAERLEPKDKGKGFVAVDELIMVSEAVOGVHNHYMWSGVEPPE
SWKPOOMPPEEPL"

CDS

BASE COUNT 245 a 345 c 357 g 245 t

ORIGIN

Alignment Scores:
Pred. No.: 8,62e-111 Length: 1192
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29,95% Indels: 0
DB: 9 Gaps: 0

US-09-784-810A-2 (1-384) x BC004112 (1-1192)

QY 270 SerGluMetPheAlaIaIaPROMetGlyArgCysAlaIaIaGlyValMetHisLeuPheTyr 289
|||||
Db 526 AGGAGATGTTGCTCCACCATGGCGCGCTGTGACGCTGGCGTCATGATCTGTTTAC 585
290 ValArgAlaGlyValSerArgAlaMetLeuLeuAlaGlyLeuPheLeuAlaMetCysGly 309
586 GTCCGGCGGAGAGTGTCTGTCGCATGCTGCGCTCTCTCTGCGCCATGAGAGAGGCG 645
QY 310 ArgHisMetGluTyrGluCysProTyrLeuValTyrValProValValAlaIaPheArgLeu 329
|||||
Db 646 AGGCAATATGGAGATGATGATCCCTTACTTGTTAGCTGCGCTGCTCCGCTTGGCTTG 705
QY 330 GluProLysAspGlyLysGlyValPheAlaValAlaGlyLeuLeuMetValSerGluAla 349
|||||
Db 706 GAGCCCAAGATGGAGAAAGGTGTTGTCAGTGTGATGGGGAATTGATGTTAGCGAGGCC 765

QY 350 ValGlnGlyGlnValHisProAsnTyrPheTyrMetValSerGlyCysValGluProPro 369
|||||
Db 766 GTGAGGCGCCAGGTGACACCCAACTACTTTCGATGTCACGGTTCGCGAGCCCGCC 825

QY 370 ProSerTyrLysProGlnGlnMetProProGluGluProLeu 384
|||||
Db 826 CCCACTGTGAAGCCCGCCACGATGTCACCGCCAGAGAGCCCTTA 870

RESULT 14
AC068145/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campitoni, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glode, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howard, J. C., Iliev, I., Johnson, R., Jones, C., Kao, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N.,
McCarthy, M., McKean, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., O'Donnell, P.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Neill, P.,
O'Neill, D., Olyar, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, L. C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Teste, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (29-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 5, 2000 this sequence version replaced gi:7960325.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/MIT Center for Genome Research
Genome Center
Center code: WIRB
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center Project name: LI0026
Center Clone name: 2531_H_7

Summary Statistics
Sequencing vector: M13; M77815; 99% of reads
Sequencing vector: Plasmid; n/a; 0.1% of reads
0.758/2534/4264/04Chemistry: Dye-terminator Big Dye; 100% of

reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152281 bases at least Q40
Consensus quality: 167680 bases at least Q30
Consensus quality: 174098 bases at least Q20
Insert size: 200000; agarose-ftp
Insert size: 178743; sum-of-configs
Quality coverage: 2.9 in Q20 bases; agarose-ftp

Quality.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 48 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 1059: contig of 1059 bp in length
1060 1159: gap of 100 bp
1160 2777: contig of 1618 bp in length
2778 2877: gap of 100 bp
2878 4283: contig of 1406 bp in length
4284 4383: gap of 100 bp
4384 5447: contig of 1064 bp in length
5448 5547: gap of 100 bp
5548 6776: contig of 1229 bp in length
6777 6876: gap of 100 bp
6877 8215: contig of 1339 bp in length
8216 8315: gap of 100 bp
8316 9961: contig of 1646 bp in length
9962 10061: gap of 100 bp
10062 12224: contig of 2163 bp in length
12225 12324: gap of 100 bp
12325 13794: contig of 1470 bp in length
13795 13894: gap of 100 bp
13895 16857: contig of 2963 bp in length
16858 16957: gap of 100 bp
16958 18266: contig of 1309 bp in length
18267 18366: gap of 100 bp
18367 20190: contig of 1824 bp in length
20191 20290: gap of 100 bp
20291 21460: contig of 1170 bp in length
21461 21560: gap of 100 bp
21561 23407: contig of 1847 bp in length
23408 23507: gap of 100 bp
23508 25455: contig of 1948 bp in length
25456 25555: gap of 100 bp
25556 27331: contig of 1776 bp in length
27332 27431: gap of 100 bp
27432 30406: contig of 2975 bp in length
30407 30506: gap of 100 bp
30507 32805: contig of 2299 bp in length
32806 32905: gap of 100 bp
32906 35659: contig of 2754 bp in length
35660 35759: gap of 100 bp
35760 37974: contig of 2215 bp in length
37975 38074: gap of 100 bp
38075 40521: contig of 2447 bp in length
40522 40621: gap of 100 bp
40622 43219: contig of 2598 bp in length
43220 43319: gap of 100 bp
43320 46161: contig of 2842 bp in length
46162 46261: gap of 100 bp
46262 48463: contig of 2202 bp in length
48464 48563: gap of 100 bp
48564 51275: contig of 2712 bp in length
51276 51375: gap of 100 bp
51376 53522: contig of 2147 bp in length
53523 53622: gap of 100 bp
53623 56955: contig of 3333 bp in length
56956 57055: gap of 100 bp
57056 59555: contig of 2500 bp in length
59556 59655: gap of 100 bp
59656 62713: contig of 3058 bp in length
62714 62813: gap of 100 bp
62814 66519: contig of 3706 bp in length
66520 66619: gap of 100 bp
66620 68930: contig of 2311 bp in length
68931 69030: gap of 100 bp
69031 72724: contig of 3694 bp in length
72725 72824: gap of 100 bp

```

```

* 72825 76092: contig of 3268 bp in length
* 76093 76192: gap of 100 bp
* 76193 80143: contig of 3951 bp in length
* 80144 80243: gap of 100 bp
* 80244 84957: contig of 4714 bp in length
* 84958 85057: gap of 100 bp
* 85058 88424: contig of 3367 bp in length
* 88425 88524: gap of 100 bp
* 88525 93428: contig of 4904 bp in length
* 93429 93528: gap of 100 bp
* 93529 97177: contig of 3649 bp in length
* 97178 97277: gap of 100 bp
* 97278 102713: contig of 5436 bp in length
* 102714 102813: gap of 100 bp
* 102814 109141: contig of 6328 bp in length
* 109142 109241: gap of 100 bp
* 109242 114513: contig of 5272 bp in length
* 114514 114613: gap of 100 bp
* 114614 124525: contig of 9912 bp in length
* 124526 124625: gap of 100 bp
* 124626 132364: contig of 7739 bp in length
* 132365 132464: gap of 100 bp
* 132465 139488: contig of 7024 bp in length
* 139489 139588: gap of 100 bp
* 139589 148734: contig of 9146 bp in length
* 148735 148834: gap of 100 bp
* 148835 159465: contig of 10631 bp in length
* 159466 159565: gap of 100 bp
* 159566 172685: contig of 13120 bp in length
* 172686 172785: gap of 100 bp
* 172786 183443: contig of 10658 bp in length.

```

FEATURES

```

source
1. 183443
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="17"
  /map="17"
  /clone="CTD-2531H7"
  /clone_1lb="CIRD Human BAC"
1. 1059
  /note="assembly-fragment"
  /note="1160..2777"
  /note="assembly-fragment"
  /note="2878..4283"
  /note="assembly-fragment"
  /note="4384..5447"
  /note="assembly-fragment"
  /note="5448..6776"
  /note="assembly-fragment"
  /note="6877..8215"
  /note="assembly-fragment"
  /note="8316..9961"
  /note="assembly-fragment"
  /note="10062..12224"
  /note="assembly-fragment"
  /note="12325..13794"
  /note="assembly-fragment"
  /note="13895..16857"
  /note="assembly-fragment"
  /note="16958..18266"
  /note="assembly-fragment"
  /note="18367..20190"
  /note="assembly-fragment"
  /note="20291..21460"
  /note="assembly-fragment"
  /note="21461..23407"
  /note="assembly-fragment"
  /note="23508..25455"
  /note="assembly-fragment"
  /note="25556..27331"
  /note="assembly-fragment"
  /note="27432..30406"
  /note="assembly-fragment"
  /note="30507..32805"
  /note="assembly-fragment"
  /note="32906..35659"
  /note="assembly-fragment"
  /note="35760..37974"
  /note="assembly-fragment"
  /note="38075..40521"
  /note="assembly-fragment"
  /note="40622..43219"
  /note="assembly-fragment"
  /note="43320..46161"
  /note="assembly-fragment"
  /note="46262..48463"
  /note="assembly-fragment"
  /note="48564..51275"
  /note="assembly-fragment"
  /note="51376..53522"
  /note="assembly-fragment"
  /note="53623..56955"
  /note="assembly-fragment"
  /note="57056..59555"
  /note="assembly-fragment"
  /note="59656..62713"
  /note="assembly-fragment"
  /note="62814..66519"
  /note="assembly-fragment"
  /note="66620..68930"
  /note="assembly-fragment"
  /note="69031..72724"
  /note="assembly-fragment"
  /note="72725..72824"
  /note="assembly-fragment"

```

Alignment Scores:

```

Pred. No.: 4,94e-109 Length: 183443
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.95% Indels: 0
DB: 2 Gaps: 0

```



```
/note="assembly_name:Contig17"
2416..4572
misc_feature
/note="assembly_name:Contig18"
4673..7009
misc_feature
/note="assembly_name:Contig19"
7110..9765
misc_feature
/note="assembly_name:Contig20"
9866..12473
misc_feature
/note="assembly_name:Contig21"
12574..16666
misc_feature
/note="assembly_name:Contig22"
16767..20013
misc_feature
/note="assembly_name:Contig23"
20114..24008
misc_feature
/note="assembly_name:Contig24"
24109..28049
misc_feature
/note="assembly_name:Contig25"
28150..31940
misc_feature
/note="assembly_name:Contig26"
32041..35923
misc_feature
/note="assembly_name:Contig27"
36024..40326
misc_feature
/note="assembly_name:Contig28"
40427..45270
misc_feature
/note="assembly_name:Contig29"
45371..48815
misc_feature
/note="assembly_name:Contig30"
48916..53904
misc_feature
/note="assembly_name:Contig31"
54005..58337
misc_feature
/note="assembly_name:Contig32"
58438..63057
misc_feature
/note="assembly_name:Contig33"
63158..67683
misc_feature
/note="assembly_name:Contig34"
67784..74319
misc_feature
/note="assembly_name:Contig35"
74420..83151
misc_feature
/note="assembly_name:Contig36"
83252..89129
misc_feature
/note="assembly_name:Contig37"
89230..98220
misc_feature
/note="assembly_name:Contig38"
98321..105559
misc_feature
/note="assembly_name:Contig39"
105660..113081
misc_feature
/note="assembly_name:Contig40"
113182..120975
misc_feature
/note="assembly_name:Contig41"
121076..132903
misc_feature
/note="assembly_name:Contig42"
133004..144491
misc_feature
/note="assembly_name:Contig43"
144592..153726
misc_feature
/note="assembly_name:Contig44"
153827..171330
misc_feature
/note="assembly_name:Contig45"
171431..191919
misc_feature
/note="assembly_name:Contig46
clone_end:r7
vector_side:right"
192020..209861
/note="assembly_name:Contig47
clone_end:SP6
vector_side:right"

BASE COUNT 52775 a 51288 c 51341 g 51413 t 3044 others
ORIGIN
```

```
Alignment Scores:
Pred. No.: 5,51e-109 Length: 209861
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 29.95% Indels: 0
DB: 2 Gaps: 0
US-09-784-810a-2 (1-384) x AC021196 (1-209861)
QY 270 SerGIuMetPheAlaAlaPrometGIYArGCysAlaAlaGlyValMetHisLeuPheTyr 289
DB 72164 AGTGAGATGTTTGGTCGACCCATGGCCGCTGTGAGCTGGCGCATGCATCATCTGTTCAC 72105
QY 290 ValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGlyLysGly 309
DB 72104 GTGCGGGCGGAGAGTCTCGTCCGATGCTGCGCCCTTCCTGGCCATGAGAGAGGC 72045
QY 310 ArgHisMetGIuTyrGIuCySProTyrLeuValTyrValProValAlaAlaPheArgLeu 329
DB 72044 AGGCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 71985
QY 330 GIuProLysAspGIuLysGlyValPheAlaValAspGIuLeuMetValSerGIuAla 349
DB 71984 GAGCCCAAGATGGGAAAGGTGTGTTCAGTGGATGGGAATTGATGTTAGCGAGGCC 71925
QY 350 ValGIuGIuGIuValHisProAsnTyrPheTrpMetValSerGIuCyValGIuProPro 369
DB 71924 GTGCAGGCGCAGGTGCACCCAAACTACTTCTGATGTGCACGCGTTCGAGAGCCCGC 71865
QY 370 ProSerTrpLysProGIuGIuMetProProProGIuGIuGIuProLeu 384
DB 71864 CCCAGCTGAAAGCCCGACAGATGCGACCGCCAGAGAGCCCTTA 71820
```

Search completed: July 12, 2003, 20:53:59
Job time : 3507 secs

THIS PAGE BLANK (USPTO)

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

3403.978 million cell updates/sec

33

32308132

11

BI005373 RC4 - HN004

Health, Mammalian

993 ATCCAGAGCCCTGTGTAGCCTCCAGCAGGCTCTGGCAACGGCTGGCAGC 1000

QY	121	AsnHs1a7yAlaG1yTyrGluGlnVal1ThrAsnGluAspPheLeuThrAspCysThrLeu	140
Db	1053	AACCATATATCTGGTATGAGCGAGTACCAATGAAAGACTCTCTGACCAACTGCACGCTA	1112
QY	141	LeuLeuGlyArgProVal1LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly	160
Db	1113	TTGGTGTGGCGCGCGGCTGCTGTACCCAGAACCTGTCTGTCTGCACACGGCTTCGGGG	1172
QY	161	LeuArgSerPheSerVal1LeuSerLeuAla1aTrrGlyPhe1a1aAspVal1aAspLeuGlu	180
Db	1173	CTGGGCGCTCTCTCTGTGTGCTACAGCTGGCTGGGCGGCTTATTGCTGATGTGGACCTAAG	1232
QY	181	SerAspLysTyrTrpArgArgLeuGlyGlnMetArgPheThrLeuGly1ThrPheLeuArgLeu	200
Db	1233	AGTGAAGAATATCGGGCGCTGGGGAGATGCGCTTACGTCTGGGCACTTCTCGGCTG	1292
QY	201	AlaAlaLeuArgThrTyrArgGlyArgLeuAla1aThrLeuProVal1aGlyArgValaGlyPhe	220
Db	1293	GCAGCGCTCGGCACCTACCGCGGCGCGACTGCTACCTCTCTAGGAAGATGGGGTTC	1352
QY	221	LysThrProAlaSerProVal1ValaGlnGlnGlyProVal1aAspAlaHisLeuVal1Pro	240
Db	1353	AAGACACCTGCTCCCTCCCGGTGGTGTCTCAGAGGGCCCGGTACATGCACACCTTGGCCA	1412
QY	241	LeuGluGluGlnVal1ProSerHisThrGlnVala1ProAspGluAspPheVal1LeuVal	260
Db	1413	CTGGAGAGACCACTGCCCTCTCACGTGACAGTGGTCCCGCAGACGACTTGTGTCTAATGTC	1472
QY	261	LeuAlaLeuLeuHisSerHisLeuAlaSerLumethPheAla1aProMetGlyArgCys	280
Db	1473	CTGGGACCTGCTCCTCCTCCACTGGGCGAGTGAAGTGTGCTGCACACCA1aGGGCGGTGT	1532
QY	281	AlaAlaGlyVala1MetHisLeuPheTyrVala1aArgAlaGlyValaSerArgAlaMetLeu	300
Db	1533	GCAGCTGGGCTCATGATGTTGTTCACGCGGGGGAGTGTCTCTGCTGCATCTGCTG	1592
QY	301	ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGlyTyrGluCysProTyrLeuVal	320
Db	1593	CGCGCTTCTCCCTGCGCCATGGAGAAGGCGAGGCAATATGGAGTATGAATGCCCTACTTGSTA	1652
QY	321	TyrVal1ProVala1a1aPheArgLeuGluProLysAspGlyLysGlyVala1PheAlaVal	340
Db	1653	TAGTGCCTGGTGTGCGCTCCGTTGTGGAGCCCAAGATGGGAAGGATGTTTGTGCAGTG	1712
QY	341	AspGlyGluLeuMetVal1SerGluAlaValaGlnGlyGlnValaHisProAsnTyrPheThr	360
Db	1713	GATGGGAAATGATGTATGCTAGCGAGGCGCTGCAGGGCCAGGTGCACCCAAACTACTTGG	1772
QY	361	MetVal1SerGlyCysValaGluProProPheSerTrrPlysProGlnGlnMetProProPro	380
Db	1773	ATGGTCAAGCGTGTGCGTGGAGCCCGCCAGCTGGAAGCCCAAGATGTCACACGCCCA	1832
QY	381	GluGluProLeu	384
Db	1833	GAGAGAGCCCTTA	1844
RESULT 2			
BM808698		1054 bp	mrna
LOCUS			linear
DEFINITION	BM808698		EST 05-MAR-2002
ACCESSION	AGNCNCOURT.65826722 NIH_MGC_41		Homo sapiens cDNA clone IMAGE:5471258
VERSION	5', mRNA sequence.		
KEYWORDS	BM808698.1	GI:19125509	
SOURCE	EST.		
ORGANISM	human.		
	Human sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1054)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D.		

Email: c9apbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM1976 row: g column: 03
 High quality sequence start: 7
 High quality sequence stop: 666.
 Location/Qualifiers

FEATURES

1..1054
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5471258"
 /clone_lib="NIH_MGC_41"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC library."

BASE COUNT 177 a 317 c 340 g 220 t
 ORIGIN

Alignment Scores:

Pred. No.: 8.38e-151 Length: 1054
 Score: 1499.50 Matches: 302
 Percent Similarity: 90.35% Conservatve: 7
 Best Local Similarity: 88.30% Mismatches: 27
 Query Match: 74.42% Indels: 6
 DB: 14 Gaps: 2

US-09-784-810a-2 (1-384) x BM808698 (1-1054)

49 PherthleuMetleuThglUARgARgAsnHsAlaArgLuleuValArgSerLuglu 68
 |||||
 3 TTCACGCTGATGTTCACTGACGCGGGAACACCGCGGAGCTGTGCGTGGAGAGG 62
 |||||
 69 LeuGIARgTRPASPAlaLeuValAlaMetSerGIyASpGLyLeuMetHsIGluValAl 88
 |||||
 63 CTGGCGCGCTGGAGAGCTCTGTGTCTCATGTCTGAGACGGGCTATGACAGAGTGGT 122
 |||||
 89 AsnGIyLeuMetGIARgProAspTRPgluThralAlaIleGlnLysProLeuCySerleu 108
 |||||
 123 AACGGGCTCATGAGACCGCTGACGAGACCGCCATCCACAAGCCCTGTGTGAGCTTC 182
 |||||
 109 ProAlaGIySerGIyASnAlaLeuAlaSerleuAsnHsTYrAlaGIyTYrLuglu 128
 |||||
 183 CCAGAGGCTCTGGGAACGCGCTGACCTTCCTGAACCATATGCTGGCTATGAGCAG 242
 |||||
 129 ValTRPAsnGIyASpLeuLeuThAsnCYSThrleuLeuLeuCyArgProValleuSer 148
 |||||
 243 GTCACCAATGAAGACCTCTGACCACTGACAGCTATGTCTGTGCGCGGCTGTCTCA 302
 |||||
 149 ProMetAsnLeuLeuSerleuHsThralaSerGIyLeuArgSerPheSerValleuSer 168
 |||||
 303 CCCATGAACCTCTCTCTGACACAGCGCTTCGGCGGCTCTCTCTGTCTGTCTGAGC 362
 |||||
 169 LeuAlaTRPGLyPheAlaAspValaAspLeuGIySerAspLysTYrArgArgleuGIy 188
 |||||
 363 CTGGCGCTGGGCTTATTGCTGATGTGAGACCTAGAGAGTGAACATATCGGCGTGGGG 422
 |||||
 189 GluMetArgPheThleuGIyThrPheleuArgleuAlaAlaLeuArgThrTYrArgGIy 208
 |||||
 423 GAGATGGCGCTTACCTCTGCGACCTTCTGCGCTGCGACGCCCTGCGACCTACCGCGGC 482
 |||||
 209 ArgLeuAlaThrLeuProValGIyArgValGIyPheLysThrProAlaSerProValVal 228
 |||||

483 CCACCTGGCTTACCTCTCTAGAGAGAGTGGGTTCCAGACACCTGCTCCCGCTGTG 542
 |||||
 229 ValGIingInGIyProValaSpAlaHsLeuValProLeuGIuGIuGIuValProSerHs 248
 |||||
 543 GTCACAGAGAGGCGCGGTATGATCAGACCTGTGTGCCACTGGAGAGAGACAGTCCCTGAC 602
 |||||
 249 TRPGLuValaValProAspGIuLysPheValleuValleuAlaLeuLeuHsSerHsLeu 268
 |||||
 603 TGGACAGTGTGCTCCGACAGAGAGCTTGTGTGATGTCCTGGCAGCTGCGACCTG 662
 |||||
 269 AlaSerGIuMetPheAlaAlaProMetGIyArgCYsAlaAlaGIyValMetHsLeuPhe 288
 |||||
 663 GGCAGTGAATTTGTCTGACACCATGGGCGGCTGTGAGCTGGCTATGATCTGTC 722
 |||||
 289 TYrValaArg-AlaGIyValaSerArgAlaMetleuLeuArgleuAlaMetGIy 308
 |||||
 723 TACGTGGCGGGGAGAGTGTCTGTGCCATGTGTGCTGCGCCCTTCTGCGCATGGAGAA 782
 |||||
 308 sGIyArgHsMetGIyTYrGIuCySProTYrleuValTYrValProValaAlaPheAr 328
 |||||
 783 GGGCAGGCTATGAGATGAAATGCCCTTACTTGGATATGTGCGGTGCTGCTTTCG 842
 |||||
 328 GluGIuProLysAspGIyLysGIyValPheAlaValaSp---GIyGIuLeuMetValSe 347
 |||||
 843 CTGGAGGCCAAGAGATGGGAAGAAGTGTGTTTCACTGAGATGGGAATGATGTGTAC 902
 |||||
 347 rGIuAlaValaGIingInGIyValaHsProAsnTYrPheThrMetValSerGIyCyVala-- 366
 |||||
 903 CGAGCGCGCTGACGAGGCGCCAGTGCACCCACTTATTTTGAGAGAGGCCCAAGCGGTGG 962
 |||||
 367 ----GluProProProSerTrpLys-ProGIingInMet-ProProProGIuLuleuProLeu 384
 |||||
 963 GTGAGACCCGCCCGCAGCTGGAACCCCGCCGAGATGCCACCGCCAAAGAACCCCTT 1022
 |||||

RESULT 3
 BM810136
 LOCUS
 DEFINITION
 AGENCOURT_6579661 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5470089
 5', mRNA sequence.
 ACCESSION
 BM810136
 VERSION
 BM810136.1 GI:19126959
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Homo sapiens
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1137)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM1976 row: f column: 10
 High quality sequence stop: 622.
 Location/Qualifiers

FEATURES

1..1137
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5470089"
 /clone_lib="NIH_MGC_41"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."

BASE COUNT 188 a 342 c 363 g 238 t 6 others
 ALIGNMENT SCORES:
 Pred. No.: 5.27e-148 Length: 1137
 Score: 1474.00 Matches: 302
 Percent Similarity: 88.51% Conservative: 6
 Best Local Similarity: 86.78% Mismatches: 25
 Query Match: 73.15% Indels: 15
 DB: 14 Gaps: 6

US-09-784-810a-2 (1-384) x BM810136 (1-1137)

49 PheprrleuMetleuThrgluArgAsnHisAlaArgluValArgSerGluGlu 68
 3 TTCACGCTGATGCTCAGAGCGGAGACCGGGGGAGTGGTGGAGAG 62
 69 LeuGlyArgTrpAspAlaLeuValAlaMetSerGlyAspGlyLeuMetHisGluVal 88
 63 CTGGGCGCGTGGAGCGCTGCTGTGCTGCTGAGAGCGGCTGATGACAGAGTGTG 122
 89 AsnGlyLeuMetGluArgProAspTrpGluThrAlaLeuGlnValProLeuGlySer 108
 123 AACGGGCTCATGAGAGCGGCTGACCTGAGAGACCGCATCAGAACCCCTGTGTAGCTC 182
 109 ProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisAlaGlyValGlyGlu 128
 183 CCAGCAGCGCTGGAGAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
 129 ValThrAsnGluAspLeuLeuThrAsnGlyLeuLeuGlyValGlyValGlyVal 148
 243 GTACACCAATGAAACCTCTCTGACCAACTGACGATTTGCTGCGCGCTGCTGCTCA 302
 149 ProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSer 168
 303 CCATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362
 169 LeuAlaTrpGlyPheLeuAlaAspValAspLeuGluSerAspGlyValArgGly 188
 363 CTGGCGTGGGCTTCATGCTGATGCTGAGACCTGAGAGAGTGAAGTGGCTGGGG 422
 189 GluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAlaLeuArgThrArgGly 208
 423 GAGATGGGCTTCACTGCGCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 482
 209 ArgLeuAlaThrLeuProValGlyArgValGlyPheLeuThrProAlaSerProVal 228
 483 CGACTGGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
 229 ValGlnGlnGlyProValAspAlaHisLeuValProLeuGluGluGlnValProSerHis 248
 543 GTCCAGCAGGCGCGGAGTGAATGCACACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 602
 249 TrpGlnValValProAspGluAspPheValLeuValLeuAlaLeuHisSerHisLeu 268
 603 TGGACAGTGGTGGCGGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
 269 AlaSerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGlyValMetHisLeuPhe 288
 663 GGCATGTGATGTTTGTGCTGACCCATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
 289 TyrValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluHis 308
 723 TACGTGGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
 309 GlyAlaGlnHisMetGluThrGluCysProTyrLeuValTyrVal-ProValAlaLaphe-A 328
 783 GGCAGCATATGATGATGATGCTTACTTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842

328 rgleuGluProGlyAspGlyLys-----GlyValPheAlaValAspGlyG 343
 843 GCTTGGAGCCAGGATGAGGAGAAAGTGTGTTTGGCANTGGATGGGAGAAATGATGAGN- 901
 343 IuleMetValSerGluAlaValGlnGlyValHisProAsnTyrPhe-TrpMet--- 361
 902 -----TTANCGAGGCGCTGGCAGGCGCAATGCACCCAACTACTCTGGAAGGCG 953
 362 ValSerGlyCysValGluProPro---ProSerTrpLysProGlnGlnMet---ProPro 379
 954 CCAAGGGGTTCGCGGAAACCCCGGCGCTTGGAAAGCCCGCCACCATATGCGCCCGC 1013
 380 ProGlu---GluProLeu 384
 1014 CCAGAAAGAAACCTTTA 1031

RESULT 4
 BC011432 2145 bp mRNA linear HTC 07-AUG-2002
 LOCUS
 DEFINITION Mus musculus, clone IMAGE:3603523, mRNA.
 ACCESSION BC011432.1 GI:15277470
 VERSION
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

FEATURES
 source
 1..2145
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="IMAGE:3603523"
 /tissue_type="mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
 /clone_id="NCI CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 Location/Qualifiers
 1..2145

BASE COUNT 447 a 581 c 620 g 497 t

ALIGNMENT SCORES: 3.72e-145 Length: 2145
 Pred. No.: 3.72e-145

/lab_host="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SportC; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 152 a 266 c 283 g 191 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,35e-132 Length: 892
 Score: 1328.50 Matches: 273
 Percent Similarity: 93.62% Conservative: 6
 Best Local Similarity: 91.61% Mismatches: 18
 Query Match: 65.98% Indels: 5
 DB: 13 Gaps: 0

US-09-784-810a-2 (1-384) x B1860351 (1-892)

OY 86 GUAUVALAAsnglyLeuMetGUArgProAspTPrpGluThrAlaIleGlnLysProLeu 105
 DB 3 GAGGTGTGAACGGGCTCATGTAGAGCGGCTGACTGGAGACCCGCAATCCAGAAAGCCCTG 62
 OY 106 CysSerLeuProAlaGlySerGlyAsnAlaLeuAlaIleSerLeuAsnHisTyrAlaGly 125
 DB 63 TGTACCTCCAGCAGCGCTTGCAACCGCTGCGACCTCTTGAACCAATTATGTGGGC 122
 OY 126 TyrGluGlnValThrAsnGluAspLeuThrAsnGlyThrLeuLeuLeuGlyCysArgPro 145
 DB 123 TATGAGCAGGTACCAATGAAAGACCTCTGACCAACTGCACGTTATGTGGCGCG 182
 OY 146 ValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSer 165
 DB 183 CTGCTGTACCATGCACTGCTCTGTGCAACAGCGCTTGCGGCGCTGCTTCTCT 242
 OY 166 ValLeuSerLeuAlaTPrpGlyPheIleAlaAspValaAspLeuGluSerAspLysTyrArg 185
 DB 243 GTCTGTACGCTGGCGGCTTCTCATGTGTGATGTGACACTAGACAGTGAAGATGCG 302
 OY 186 ArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAlaLeuArgThr 205
 DB 303 CGCTGTGGGGAGATGGGCTTCACTGTGGCACCTTCTGCTGGACCCCTGGCGACC 362
 OY 206 TyrArgGlyArgLeuAlaThrLeuProValaGlyArgValaGlyPheLysThrProLaser 225
 DB 363 TACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
 OY 226 ProValaValaGlnGlnGlyProValaAspAlaHisLeuValaProLeuGluGlnVala 245
 DB 423 CCCGTGTGTGCGCAGCGGCGCGGTAGATGCACACTTGTGCACTGAGAGAGCCAGTG 482
 OY 246 ProSerHisTPrpGlnValaValaProAspGluAspPheValLeuValaLeuAlaLeuHis 265
 DB 483 CCCGTGTGTGCGCAGCGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
 OY 266 SerHisLeuAlaSerGluMetPheAlaAlaProMetCylArgCysAlaAlaGlyValaMet 285
 DB 543 TCGCACCTGGGCACTGAGATGTT-GCTGCAACCATGGCGCGCTGTGAGCTGCGCTCATG 601
 OY 286 HisLeuPheTyrValaArgAlaGlyValaSerArgAlaMetLeuLeuArgLeuPheLeuAla 305
 DB 602 CATCTGTCTACGTGCGGCGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
 OY 306 MetGluLysGlyArgHisMetGlyTyrGluCysProTyrLeuValaTyrValaProValaVal 325
 DB 662 ATGGAAGAAGGCGCATTTGGATGATGAATGCCCTACTGCTGATGATGCGCGGTGTC 721
 OY 326 AlaPheArgLeuGlnProLysAspGly-LysGlyValaPheAlaValaAspGlyLeuLeu 345
 DB 722 GCCTTCGCTTGGAGCCCAAGATGCGACAGGTGTGTTGCTGAGTGGCGCAAGATGAT 781
 OY 345 tValSerGluAlaValaGlnGlyGlnValaHisProAspTyrPheThrPheValaSerGlyC 365

DB 782 GGTACGAGGCGGCTGACAGGCCAGTGCAGCAAA-TACTTCTGATGTACAGCGGTG 840

OY 365 ValGluProProSerTPrpLysPProGlnGlnMetProProGluGlu 382
 DB 841 CTG-GAGCCCGCCGCGCAGC-TGGAAGGCCAGCAGATGCCAGCGCAGAAAG 890

RESULT 6
 LOCUS BM916526 1100 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT_6641989 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482958
 5', mRNA sequence.
 ACCESSION BM916526
 VERSION BM916526.1 GI:19366905
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 1100)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.jhmi.gov
 Plate: LNCM2009 row: n column: 15
 High quality sequence stop: 625.
 Location/Qualifiers

FEATURES
 source
 1..1100
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5482958"
 /clone_1ib="NIH_MGC_41"
 /tissue="type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCGAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using zap-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

BASE COUNT 187 a 333 c 348 g 222 t 10 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.35e-122 Length: 1100
 Score: 1238.00 Matches: 275
 Percent Similarity: 81.18% Conservative: 14
 Best Local Similarity: 77.25% Mismatches: 46
 Query Match: 61.44% Indels: 22
 DB: 14 Gaps: 9

US-09-784-810a-2 (1-384) x BM916526 (1-1100)

OY 49 PheThrLeuMetLeuThrGluArgArgAsnHisAlaArgGluLeuValaArgSerGluGlu 68
 DB 3 TTACAGCGTGAAGCTCAGTGAAGCGGCGGAGACCGCGGAGCTGTGCGGTGAGAG 62
 OY 69 LeuGlyArgTPrpAspAlaLeuValaValaMetSerGlyAspGlyLeuMetHisGluVala 88
 DB 63 CTGGGCGCGTGGAGCGCTCTGTGCTCATGTCTGTGAGACGCGCGTGAAGCAGAGTGTG 122
 OY 89 AsnGlyLeuMetGluArgProAspTPrpLThrAlaIleGlnLysProLeuCysSerLeu 108
 DB 123 AACGGCTCATGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182

109 ProAlaGlySerGlyAsnAlaLeuAlaLeuSerLeuAsnHisTyrAlaGlyTyrGlu 128
 183 CCAGACAGCCTCTGGCAACGCGCTGGACCTCTTGAACCATTAATGCTGATGAGCAG 242
 129 ValThrAsnGluAspLeuLeuThrAsnCysThrLeuLeuLeuCysArgProValLeuSer 148
 243 GTCACACATGAAGACCTCTGACCAACATGACGATATGCTGCGCGCGCTGCTGCA 302
 149 ProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSer 168
 303 CCATGAACCTGCTCTCTCTGACACGCGCTTGGGGGCTCCCTCTCTCTGCTGCTGAC 362
 169 LeuAlaTrpGlyPheLeuAlaAspValAspLeuGluSerAspLysTyrArgArgLeu 188
 363 CTGGCTTGGGCTTATGCTGATGAGTGGACCTAGAGAGTGAAGATATGCGGCTGGGG 422
 189 GluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAlaLeuArgThrTyrArgGly 208
 423 GAGATGGCCTTACACTCTGGCACCTTCTGGCTGGGACGCTGGCACCTACCGGCGC 482
 209 ArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrProAlaSerProValVal 228
 483 CGACTGGCCTACCTCCCTGAGGAAGAGTGGGTTCCAGACACCTGCCCTCCCTGCTG 542
 229 ValGlnGlnGlyProValAspAlaHisLeuValProLeuGluGlnGlnValProSerHis 248
 543 GTCCACGACGGCGCGGATGATGACACCTGTGTGCTGAGAGAGACCATGCGCTCTAC 602
 249 TrpGlnValValProAspGluAspPheValLeuValLeuAlaLeuHisSerHisLeu 268
 603 TGGACAGTGTGCGCGAGAGAGACTTGTGCTAGTCTGGGACCTGCTGACCTGCGACCTG 662
 269 --AlaSerGluMetPheAlaAlaProMetGly-ArgCysAlaAlaGlyValMetHisLeu 287
 663 GGCAGATGAGATGTTCTCTGACCAATGGCGCCCTGTGACGCTGATGATGATGCTG 722
 288 PheTyr--ValArgAlaGly--ValSerArgAlaMetLeuAlaArgLeuPheLeu-Al 305
 723 TTCACGCGCNCNNGGCGGGAAGTGTCTCGCTGACATGGCTGCTGCTCTTCTGCGNCC 782
 305 aMetGluLysGlyArgHisMetGluTyrGlu-----CysProTyrLeu-ValTyrVal 322
 783 CTGGGAAAAAGGCGCGCCATT--TGGGAAGTATGATGAGCGCCCTAATGAGATATGG 839
 322 alProValValAlaPheArgLeuGluProLysAsp--GlyLysGlyValAlaPheAlaVal 341
 840 GCGCGGTGGCGGCTTCCGCTTGGAAACCCAGGAGTGGGAAAGGGTGTTCCTGCAAGG 899
 341 sp--GlyGluLeu-----MetValSerGluAlaValGlnGlyGlnValHisProAsnT 358
 900 GAATGGGGGAATTAATGATTACCAAGCCCGGCGGACGAGGCGCCCAAGGGGCAAT 959
 358 yPheTyrMetValSer-----GlyCysValGluProProProSerTyrLys- 373
 960 TACCTTTTGGATTGGCCCAAGGGGTTTGGGGGGAACCCCGGCGCCCATCTTGGGA-A 1018
 374 -----ProGlnGlnMetProProProGluGluProLeu 384
 1019 CAGCCCCNCAATATGTCCTCCCTCCCAAAAAGCCCTT 1056
 RESULT 7
 B0647377 888 bp mRNA linear EST 15-JUL-2002
 LOCUS B0647377
 DEFINITION AGENCOURT_8414975 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6271071
 5', mRNA sequence.
 ACCESSION B0647377
 VERSION B0647377.1 GI:21771549
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 888)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: CGAP (Starford)
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L10M2449 row: n column: 08
 High quality sequence start: 6
 High quality sequence stop: 668.
 Location/Qualifiers
 1..888
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6272071"
 /clone_11b="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; XhoI sites using the following 5' adaptor:
 GGACAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."
 BASE COUNT 137 a 274 c 284 g 193 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8,466-121 Length: 888
 Score: 1220.00 Matches: 237
 Percent Similarity: 96.408 Conservative: 4
 Best Local Similarity: 94.808 Mismatches: 7
 Query Match: 60.55% Indels: 2
 DB: 14 Gaps: 0
 US-09-784-810A-2 (1-384) x B0647377 (1-888)
 137 AsnCysThrLeuLeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHis 156
 31 CACTGCACGCTATGCTGTCGCGCGGCTGCTGTCACCATGACCTGCTCTGAC 90
 157 ThrAlaSerGlyLeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAsp 176
 91 ACGGCTTGGGGGCTGCGCTCTCTCTGCTGCTGACCTGGGCTGCTCAATGCTCAT 150
 177 ValAspLeuGluSerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThr 196
 131 GTGACCTAGAGAGAGAGATGCGGCTGCGGAGATGCGCTTCACTCTGGGACAC 210
 197 PheLeuArgLeuAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGly 216
 211 TTCTGCTGCTGGAGCGCTGCGCACATGACCGCGCGAGCTGCTACCTCTGATAGA 270
 217 ArgValGlyPheLysThrProAlaSerProValValValGlnGlnGlyProValAspAla 236
 271 AGAGTGGGTTCACAAACACCTGCTCCCGCTTGTGCTCACAGAGCGCGGTATGATCA 330
 237 HisLeuValProLeuGluGlnGlnValProSerHisTyrPglValValValProAspGlu 256
 331 CACTTGTGCGCACCTGAGAGAGCACTGCTCTCACTGACAGTGTGCTCCGACAGAGAC 390
 257 PheValLeuValLeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaPro 276
 391 TTGTGCTAGTCTGCGACCTGCTGACATGCGACCTGGGCAATGAGATGTTGCTGACACC 450

Mon Jul 14 09:16:48 2003

us-09-784-810a-2.rst

Page 8

QY	277	MetGlyArgGlyAlaIaIaGlyAlaMetHisLeuPheTyrValAlaAlaGlyValSerArg	296
Db	451	ATGGGCGCGCTGTGCACCTGGCTTCATGCTGTTCTACGTGGGGGGAGTCTTCGT	510
QY	297	AlaMetLeuLeuArgLeuPheLeuAlaMetIuIuGlyAlaArgHisMetIuTyrGlyCys	316
Db	511	GGCATTGGCGTGGGCGCTCTTCTCGGCATTGGAGAGGCGCAGCATATGATATATGC	570
QY	317	ProTyrLeuValTyrValProValValAlaPheArgLeuGluProTyrAspGlyTyrGly	336
Db	571	CCCACTCTGGATATATGTGCGCTGGCTCCCTCCGCTGTGGAGCCAGAGATGGGAAAGT	630
QY	337	ValPheAlaValAspGlyIuLeuMetValSerAlaValGlnGlyValHisPro	356
Db	631	GGTTTGCAAGGATGGGAAATTGATTGACAGAGCCGTGAGGCCAGAGTACCCA	690
QY	357	AsnTyrPheIlePheMetValSer--GlyCysValGluProPro--ProSerTyrPylSerGlnG	376
Db	691	AACATCACTTCGTGATGTACAGCGGATGGCTGGTGGAGGCCCGCCGCTGGAGAGCCCAAGC	750
QY	376	IuMetProProProGluGluProIuLeu	384
Db	751	AATATGCACACGCGAGAGAGCCCTTAA	776

RESULT 8				
BG680521	BG680521	785 bp	mRNA	linear EST 01-MAY-2001
LOCUS	G02829334F1 NC1_CGAP_Skn4	Homo sapiens	cDNA clone IMAGE:4753187 5' /	
DEFINITION	mRNA sequence.			
ACCESSION	BG680521			
VERSION	BG680521.1	GI:13911918		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

REFERENCE AUTHORS TITLE JOURNAL	COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo 1 (bases 1 to 785)	
NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	Contact: Robert Strausberg, Ph.D.

Email: cgapds-remail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA library preparation: Life Technologies, Inc.
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LNA10612 row: g column: 12
 High quality sequence: spot: 783.

FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4753187"
/clone_id="NCL_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCL_CGAP Library."
BASE COUNT
ORIGIN
130 a      231 c      250 g      174 t

```

Alignment Scores:	
Pred. No.:	1.02e-120
Score:	1218.50
Percent Similarity:	95.00%
Best Local Similarity:	93.85%
Query Match:	60.47%
Length:	785
Matches:	244
Conservative:	3
Mismatches:	12
Indels:	3

```
DB:      12
Gaps:    0
```

US-09-784-810A-2 (1-384) x BG680521 (1-785)

125 G1YTYRGIInValThrasnGIuAspleuLeuThrasncystThrLeuLeuLeucysArg 144

Db 3 GGCTATGAGCAGGTCACCAATGAGACCTCTTGACCACATCTGCCTATTGCTGTGCCG 62

143 FLOWALPUBEL FLOWECASINDEUBUBEL DEUNIS IININUBELVI DEUNIGUBELNE 104

[illegible]

THE UNIVERSITY OF CHICAGO

183 CGGCGTCTGGGAGATGCGCTTCACTCTGGGCACTTCTGCGCTGTCAGGCTCTGCGG
242

205 ThrTVrARqG]VArqLeuAlaTThrLeuProVa]G]VArqVa]G]VpPhelVstThrProAla 224

Db 243 ACCTACCGGGCGACTGGCCTACCTCCCTGTAGGAAGAGTGGGTCCAGACACCTGCC 302

225 SerProValValGlnGlnGlyProValAspAlaHisLeuValProLeuGlnGln 244

Db 303 TCCCCCGTTGTGTCACAGCAGGGCCCGGTAGATGCACACCTTGTGCCACTGGAGGAGCCA 362

245 ValProSerHisTrpGlnValValProAspGluAspPheValLeuValLeuAlaLeuLeu 264

D6 363 GTGCCCTCTTACATGGACAGTGGTGCCCGACGAGGACTTTGTGCTAAGTCCCTGGCAGTGCIG 422

265 hissehlisledualsberglumelcnealadalptometgluayrgcysaladadglval 284

20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050

[illegible]

1000

|||||

325 ValAlaPheAsnGlyValValPheAlaValAsnGlyValIle 344

Db 603 GTCCCTTCGCTTTCGAAGCCCAAGGATGGGAAAGGTGTGTTTCAGTGGATGGGAAAT 662

QY 344 umetvalserglualavalnglglvalhlsproasntyrphetrpmetvalsergl 364

Db 663 GATGGTTAGCGAGGCCCTGCAGGGGCCAGGTGCACC - AAATACTTCTGATGTCAGCG 720

364 YCYSVALGLUPROPROSERTIPLYSPROGLINMETPROPROGLINLUPRO 363

Db 721 TTGCGTGAGCCCCCGCCAGC-TGGAAGCCCCCAGCAGATGCCACC CGCAGAGAGGCCCT 777

RESULT 9
BT648186

LOCUS	BI040100	242 bp	INRA	LINE1	ESI 12 SEP 2001
DEFINITION	603279314F1 NCI CGAP Mam3 Mus musculus cDNA clone IMAGE:5319675 5'				

Accession BI648186

KEYWORDS EST.

ORGANISM Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
EDITOR National Institutes of Health, National Cancer Center (NCC)

JOURNAL
unpublished (1999)
Contact: Robert Strausberg ph D

QY 48 SerPheThrLeuMetLeuThrGluArgArgAsnHisAlaArgGluLeuValArgSerGlu 67
 DB 2 TCCTTTCACGGTGTACACTGAGCGGGGAACACCGCGAGAGTGGTGGCGGAG 61
 QY 68 GluLeuGlyArgTrpAspAlaLeuValMetSerGlyAspGlyLeuMetHisGluVal 87
 DB 62 GAGCTGGCGCGTGGAGCGCTGTGGTGTATGTCGTGAGACGGGCTGATGACAGAGGTG 121
 QY 88 ValAsnGlyLeuMetGluArgProAspTrpGluThrAlaLeuGlnProLeuGlySer 107
 DB 122 GTGACGGGCTCAGAGGAGCGCTGACGAGAGCGCGCTACCAAGAGCCCGTGTAC 181
 QY 108 LeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisTyrAlaGlyTyrGlu 127
 DB 182 CTCACACAGAGCTCTGGCAACCGCTGGACCTTCTTGAACATATGCTGCTATGAG 241
 QY 128 GlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeuLeuLeuGlySerProValLeu 147
 DB 242 CAGGTACCAATGAAAGACCTCTGACCAACTGACAGCGTATTGCTGTGGCGCGGCTGCT 301
 QY 148 SerProMetAsnLeuSerLeuThrAlaSerGlyLeuArgSerPheSerValLeu 167
 DB 302 TCACCAAGAGCTGCTGTCTGTGACAGCGCTGGCGCGCTGCTGTCTGTGCTC 361
 QY 168 SerLeuAlaTrpGlyPheIleAlaAspValAspLeuGluSerAspTyrArgArgLe 187
 DB 362 AGCTGTGGCTGGGGCTTCAATGCTGATGTGAGACCTAGAGAGAGTATCGGCTCT 421
 QY 187 uGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaLeuArgThrTyr 207
 DB 422 GGGGAGAGATCGCTTCACCTCTGCGACCTTCTGCTGTGGACCGCTGGACCTACCG 481
 QY 207 GGLYArgLeuAlaThrLeuProValGlyArgValGlyPheLeuThrProAlaSerPro-- 226
 DB 482 CGGCGAGCTGCTACCTCTCTGTAGAGAGTGGGTTCACAGACACTGCTACCTCCACC 541
 QY 227 ValValValGlnGlnGlyProValAspAlaHisLeuVal-ProLeuGlnGlnGlnVal-P 246
 DB 542 GTTGTGTCCAGAGAGGCGCCGGTAGATGACACCTGTGTGACACTGAGAGAGCACTGGC 601
 QY 246 roSerHisTrpGlnValValProAspGluAspPhe-ValLeuValLeuAlaLeuHis 265
 DB 602 CATCTCACTGGACAGTGGTGGCCGAGAGACTATGTGCTAGTCCGAGCAGTGGCAGC 661
 QY 266 SerHisLeuAlaSer---GluMetPheAlaAlaProMetGlyArg-CysAla-AlaGly 284
 DB 662 TCGACCTCGGACAGTGGCAGACTGTAGCTGCACCGCATGGCGGATGTCCAGCTGGCG 721
 QY 284 aLMethHisLeuPheTyrValArgAlaGlyValSerArgAla--MetLeuLeuArgLeu 303
 DB 722 TCATGCAATCTGCTCAGTGGCGGAGGGAATGTATGCTGCTGACCTGACCGCTCT 781
 QY 303 heLeuAlaMet 306
 DB 782 TCCTGGCCATG 792
 RESULT 11
 LOCUS BM006005 875 bp mRNA linear EST 30-OCT-2001
 DEFINITION 603613736F1 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:5433632 5',
 mRNA sequence.
 ACCESSION BM006005
 VERSION BM006005.1 GI:16520359
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH_MGC http://mgc.ncl.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: L10M1909 row: 9 column: 09
 High quality sequence stop: 714.
 Location/Qualifiers

FEATURES

source

1..875

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="5433632"

/clone_lib="NIH_MGC_100"

/tissue_type="hepatocellular carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: liver; Vector: pORF7; Site: 1; XhoI; Site: 2; EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Size-selected by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using Zap-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT

159 a 252 c 272 g 192 t
 ORIGIN

Alignment Scores:

Pred. No.: 2,72e-105 Length: 875
 Score: 1076.00 Matches: 222
 Percent Similarity: 91.97% Conservative: 7
 Best Local Similarity: 89.16% Mismatches: 13
 Query Match: 53.40% Indels: 7
 DB: 13 Gaps: 2

US-09-784-810A-2 (1-384) x BM006005 (1-875)

QY 132 GluAspLeuLeuThrAsnGlyThrLeuLeuGlySerPheSerValLeuSerLeuAlaTrp 151
 DB 2 GAAGACCTCTGACCAACTGACAGCTATGCTGCTGCCGCGCTGTGCTACCACTGAGC 61
 QY 152 LeuLeuSerLeuHisTrpAlaSerGlyLeuArgSerPheSerValLeuSerLeuAlaTrp 171
 DB 62 CTGCTGCTCTGACACAGGCTTGGGGCTGGCGCTCTCTGTGCTGCTGCTGCTGCTGCTG 121
 QY 172 GlyPheIleAlaAspValAspLeuGluSerAspTyrArgArgLeuGlyGluMetArg 191
 DB 122 GCTTCATGCTGTGTGTGGACCTGAGAGTGAAGTATGCGCTGTGGGGAGATGGC 181
 QY 192 PheThrLeuGlyThrPheLeuArgLeuAlaLeuAlaLeuArgThrTyrArgGlyArgLeuAla 211
 DB 182 TTCACCTCGGACACTTCTGCTGTGGCAGCGCTGCGACACTACCGCGCGCGCAGTGGCC 241
 QY 212 ThrLeuProValGlyArgValGlyPheLeuThrProAlaSerProValValGlnGln 231
 DB 242 TACCTCCCTGTAGAGAGAGTGGTTCAGAGACACTGCTGCTGCTGCTGCTGCTGCTGCTG 301
 QY 232 GlyProValAspAlaHisLeuValProLeuGlnGlnGlnValProSerHisTrpGlnVal 251
 DB 302 GGCCCGGTAGTACACCTGTGTGCGACGTGGAGGAGCAGGCGCTTCCACTGAGACAGT 361
 QY 252 ValProAspGluAspPheValLeuValLeuAlaLeuLeuHisSerHisLeuAlaSerGlu 271
 DB 362 GTGCCGACGAGGAGCTTGTGTGCTGCTGCGACACTGCTGCTGCGACCTGGGCGAGT 421
 QY 272 MetPheAlaAlaProMetGlyArgGlyAlaAlaGlyValMetHisLeuPheTyrValArg 291
 DB 422 ATGTTGCTGACCACTGAGCGCGCTGTGAGCTGGCGCTGATGCACTCTTCTACGTGGCG 481

OY 292 AAGlyValSerArgAlaMetLeuLeuArgPheLeuAlaMetGlyLysGlyArgHis 311
 DB 482 GCGGAGTGTCTCGCATGCTGCTGCTCTTCCTCCGCGCATGAGAAAGGAGGATG 541
 OY 312 MetGlu-TyrGlyCysProTyrLeuValTyrValProVal-ValAlaPheArgLeuGluP 331
 DB 542 ATGGACGATGTAAGTCCCTACTGCTATGATGTCCTGATGCTCCGCTGAGAGC 601
 OY 331 ROLYASPGlyLysGlyValPheAlaValAspGlyLysLeuMetValSer---GluAlaVal 350
 DB 602 CCAAGATGGAGAAAGGTGTGTTCAGATGATGGGAAATGATGCTTCACCGAGAGACG 661
 OY 350 aGIngly-GlnValHisPro-AsnTyrPheTyrPheValSer-GlyCysValGluProp 369
 DB 662 TGCAGGGCCGACGATGCGACCAAACTACTTCTGTGATGTCACGCGGTTCGCTGAGACCA 721
 OY 369 ro---ProSerTyrLys 373
 DB 722 CCAGACCAAGCTGGAAA 738
 RESULT 12
 BE891653
 LOCUS BE891653 659 bp mRNA linear EST 20-OCT-2000
 DEFINITION 601434552P1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3919548 5',
 mRNA sequence.
 ACCESSION BE891653
 VERSION BE891653.1 GI:10351191
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 659)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC/DOCP/DRP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM9749 row: h column: 13
 High quality sequence stop: 585.
 FEATURES
 source
 Location/Qualifiers
 1..659
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="3919548"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1. NotI;
 Site: 2. SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 BASE COUNT 107 a 200 c 202 g 150 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,85e-95 Length: 659
 Score: 983.00 Matches: 199
 Percent Similarity: 94.47% Conservatve: 6
 Best Local Similarity: 91.71% Mismatches: 10
 Query Match: 48.78% Indels: 3
 DB: 12 Gaps: 0
 US-09-784-810a-2 (1-384) x BE891653 (1-659)
 OY 124 AAGlyTyrGluGlnValThrAsn-GluAspLeuLeuThrAsnCysThrLeuLeuLeuLeu 143

DB 12 GCTGGCTATGAGCAGGTGACACCACTGAAGACCTCTCTGACCAACTGACGATATGCTGTG 71
 OY 143 SarGProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSe 163
 DB 72 CCGCGGGCTGTGTCACCACTGAACCTCTCTCTGACACAGGCTTCGGGGCTGCGCCT 131
 OY 163 RPhSerValLeuSerLeuAlaATPGlyPheIleAlaAspValAspLeuGluSerAspLy 183
 DB 132 CTTCCTGTGCTACAGCTGCGCTGGGCTTCAATGCTGATGTGGACCTAGAGAGTGA 191
 OY 183 STYrArgArgLeuGlyGlyLysMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAla 203
 DB 192 GTATCGGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCCTCGCTGCGGACGCT 251
 OY 203 uArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrPr 223
 DB 252 GCGCACCTACCGCGCGCGCTGCTTACCTCTCTGTAGAGAGTGGGCTTCACAGACAC 311
 OY 223 OAlaSerProValValAlaGInGlyProValAspAlaHisLeuValProLeuGluG 243
 DB 312 TGCCCTCCCCGCTGTGTGTCACAGAGGCGCCGCTGATGACACCTTGTGCCACTGGAGGA 371
 OY 243 uGlnValProSerHisTyrGlnValAlaProAspGluAspPheValLeuValAlaAla 263
 DB 372 GCGAGTGGCTCTCTGCTGACAGTGTGTCGCGACAGACACTTGTGTAGTCTGCGCACT 431
 OY 263 uLeuHisSerHisLeuAlaSerGlyLysMetPheAlaAlaProMetGlyArgValAlaAla 283
 DB 432 GCTGACACGACCGACCTGCGAGTATGTTGCTGACCCATGGGCGCTGTGACGCTGG 491
 OY 283 YValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuP 303
 DB 492 CGTATGATGTGTCTTACGTGCGGGCGGAGCTGTGTGCAAGCTGATGCGCTCTT 551
 OY 303 eLeuAlaMetGlyLysGlyArgHisMetGlyTyrGlyCysProTyrLeuVal-TyrValP 323
 DB 552 CCTGCGCATGAGAAAGGCGACGATATGAGATGAATGACCCCTATTGTGATACGTCG 611
 OY 323 roValValAlaPheArgLeuGluProLysAspGlyLysGlyValAlaPhe 338
 DB 612 CCGTGATCATCCTTCGCTTGGAGAC-CAAGATGGGAAAGGTGTGAT 657
 RESULT 13
 AV694791 671 bp mRNA linear EST 16-JAN-2002
 LOCUS AV694791 GKC Homo sapiens cDNA clone GKC8E05 5', mRNA sequence.
 DEFINITION AV694791
 ACCESSION AV694791
 VERSION AV694791.1 GI:10296654
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 671)
 AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
 Xiao,H., Gu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
 Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X.,
 Hu,G., Gu,J., Chen,X. and Han,Z.
 Insight into hepatocellular carcinogenesis at transcriptome level
 by comparing gene expression profiles of hepatocellular carcinoma
 with those of corresponding noncancerous liver
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 FEATURES
 Location/Qualifiers


```

Db 369 AG-GAGCTGTGTGTGACAGAGAGTTGGGCTCAGTGGAGCCCTGGCAGTATGTCGGT 427
Oy 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTyrGluThrAla 100
Db 428 GATGCTCATGATGATGAGTGTGATGGCTAATGGAACGCCGCGACTGGAGAGCTGCC 487
Oy 101 IleGlnLysProLeuCySerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu 120
Db 488 ATCCAGAAACCCCTGTGTAGCTCCCTGGAGGCTCCGGCAATGCCCTGGACCTCTGTG 547
Oy 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeu 140
Db 548 AACCATATGCTGTGTAGACAGAGCTGATGAAAGACCTGCTCATCACTGACGACCTG 607
Oy 141 LeuLeuCySerProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
Db 608 CTGTGTGTGCGCCGCGGCTGTCTACCCATGAACCTGCTGCTGCTGACACAGCTGTGG 667
Oy 161 LeuArgSerPheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeuGlu 180
Db 668 CTGGGCTCATATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 727
Oy 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
Db 728 AGTGAAGATGACAGCGCTTGGGGAGATGCTTTCACAGTGGGACCTTCTTGGCTTA 787
Oy 201 AlaAlaLeuArgTyrArgGlyArgLeuAlaThrLeuProValArgValGlyPhe 220
Db 788 GCACCTGCTCATATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 847
Oy 221 LysThrProAlaSerProValAlaValAlaGlnGlnGlyProValAspAlaHisLeuValPro 240
Db 848 TAAGAGACCGGCTCT--ACACTGCTGCAAGAAAGGCCCGCTGACAA-CACCTTCTCT 903
Oy 241 LeuGlu-GluGlnValProSerHisTyrGlnVal 251
Db 904 CTGGAGGAGCGCAGTGTCTCTCATTTGAGACTGTG 937

RESULT 15
LOCUS B0675531 999 bp mRNA linear EST 15-JUL-2002
DEFINITION AGCCOURL_8196565 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6256195
5', mRNA sequence.
B0675531
ACCESSION B0675531.1 GI:21786365
VERSION EST.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 999)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c99pbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
Plate: L1CM2408 row: h column: 20
High quality sequence stop: 538.
Location/Qualifiers
1..999
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6256195"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

```

/note="Organ: salivary gland. Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI. cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH-MGC Library."
BASE COUNT 175 a 337 c 315 g 171 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,066-92 Length: 999
Score: 960.00 Matches: 189
Percent Similarity: 95.98% Conservative: 2
Best Local Similarity: 94.97% Mismatches: 8
Query Match: 47.64% Indels: 0
Gaps: 0
US-09-784-810a-2 (1-384) x B0675531 (1-999)
Oy 42 LeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAlaArg 61
Db 3 TTGGCTGAGCGCTGAATCTCTTCACTGATGCTCAGTGGAGCGGCAACCGCGCG 62
Oy 62 GluLeuValArgSerGluLeuGluLeuArgTyrPheAlaLeuValMetSerGlyAsp 81
Db 63 GAGCTGTGCGGCTCGAGAGAGCTGGCGCGCTGGACGCTGTGTGCTCATGCTGAGAC 122
Oy 82 GlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTyrGluThrAlaIle 101
Db 123 GGGCTGATGACAGAGAGTGTGTGAAGGGCTCATGAGAGCGGCTGAGTGGAGACCGCATC 182
Oy 102 GlnLysProLeuCySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsn 121
Db 183 CAGAAAGCCCTGTGTAGCCTCCACAGAGCTGTGGCAAGCGCTTCTTGTGAAC 242
Oy 122 HisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeu 141
Db 243 CATTAATGCTGCTATGAGACAGTCACTCAATGAACCTCTCTACCAACTGACCGCTATTG 302
Oy 142 LeuCyArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeu 161
Db 303 CTGTGCGCGGCGCTGTGACCACTGATGAAACCTGCTGTCTGTCACACGCGCTTGGGGCTG 362
Oy 162 ArgSerPheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeuGluSer 181
Db 363 CGCCCTTCTCTGTGCTCTGAGCTGCGCTGGGCTTCTCATTTGATGTGAGCTAGAGAT 422
Oy 182 AspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeuAla 201
Db 423 GAGAAATATGCGCGCTGTGGGAGATGCGCTTCACTGTGGACACTTCTGCGTGTGGCA 482
Oy 202 AlaLeuArgTyrArgGlyArgLeuAlaThrLeuProValAlaGlyArgValGlyPheLys 221
Db 483 GCGCTGCGGAGCTACCGCGGCGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
Oy 222 ThrProAlaSerProValAlaValAlaGlnGlnGlyProValAspAlaHisLeuValPro 240
Db 543 ACACCTGCTCTCCCGCTGTGTGTCCACAGGCGCGGAGATGACACACTCGGGGCC 599

```

Search completed: July 12, 2003, 19:45:22
 Job time : 1835 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:12:38 ; Search time 216 Seconds
(without alignments)
2807.103 Million cell updates/sec

Title: US-09-784-810A-2
Perfect score: 2015
Sequence: 1 MDPAGPGRVGLPPRCRVLYL.....CVPPEPSMKPQMPPEPEPL 384

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1105431 seqs, 789497651 residues
Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlh
-O/cgnt2_1/USPRO.spool/US09784810/Unat_11072003_103442_26587/app_query.fasta.1.583
-DB-Published.Applications_USA-OFMT-fastap-SUFFIT-rnp-MINMATCH=0.1
-LOOPCL=0-LOOEXT=0-UNITS-bits-START=1-END=1-MATRIX-blosum62
-TRANS-numa40.cdi-LIST=45-DOCALIGN=200-THR_SCORE=pct-THR_MAX=100
-THR_MIN=0-ALIGN=15-MODE-LOCAL-OUTFMT-pio-NORM-ext-HEAPSIZE=500-MINLEN=0
-MAXLEN=200000000-USER-US09784810@CGN1.1.57@unat.11072003.103442.26587
-NCPU=6-ICPU=3-NO-MAP-LARGOQUERY-NEG_SCORES=0-WAIT-DSPLOCK=100
-LONGLOC-DEV-TIMEOUT=120-WARN-TIMEOUT=30-THREADS=1-XGAPOP=10-XGAPEXT=0.5
-FGAPOP=6-FGAPEXT=7-YGAPOP=10-YGAPEXT=0.5-DELOP=6-DELEXT=7

Database:

Published Applications NA:*

- 1: /cgnt2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
- 2: /cgnt2_6/ptodata/1/pubpna/PCT_NEM_PUB.seq:*
- 3: /cgnt2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
- 4: /cgnt2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgnt2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 6: /cgnt2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 7: /cgnt2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
- 8: /cgnt2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgnt2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
- 10: /cgnt2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgnt2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
- 12: /cgnt2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgnt2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
- 14: /cgnt2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2015	100.0	1600	US-09-784-810A-1
2	1962	97.4	1155	US-09-970-516-1
3	1927	95.6	1533	US-09-023-282-90
4	1594.5	79.1	1759	US-09-784-810A-3

5	1587.5	78.8	1149	US-09-970-516-5	Sequence 5, Appl1
6	875	43.4	2698	US-09-817-676A-11	Sequence 11, Appl1
7	856.5	42.5	1857	US-09-970-516-3	Sequence 3, Appl1
8	856.5	42.5	2380	US-09-817-676A-13	Sequence 13, Appl1
9	481	23.9	480	US-09-783-590-9248	Sequence 9248, Ap
10	365	18.1	296	US-10-015-219-658	Sequence 658, App
11	365	18.1	296	US-09-777-564-658	Sequence 658, App
12	360	17.9	199	US-09-796-692-2905	Sequence 2905, Ap
13	360	17.9	199	US-10-040-862-2905	Sequence 2905, Ap
14	268.5	13.3	1614	US-09-969-896-9	Sequence 9, Appl1
15	268.5	13.3	4413	US-09-969-896-16	Sequence 16, Appl1
16	258	12.7	1840	US-09-784-810A-5	Sequence 5, Appl1
17	228	11.3	394	US-09-954-456-1756	Sequence 1756, Ap
18	223	11.1	979	US-09-969-896-1	Sequence 796, App
19	198	9.8	2462	US-10-037-270-796	Sequence 42, Appl1
20	158.5	7.9	3758	US-09-948-820-42	Sequence 16, Appl1
21	158.5	7.9	4129	US-09-948-820-16	Sequence 2836, Ap
22	151	7.5	1095	US-09-738-626-2836	Sequence 98, Appl1
23	151	7.5	1095	US-09-738-626-1	Sequence 1, Appl1
24	147.5	7.3	16870	US-09-070-927A-98	Sequence 98, Appl1
25	142.5	7.1	9025608	US-10-156-761-1	Sequence 1, Appl1
26	137	6.8	957	US-10-156-761-2261	Sequence 2261, Ap
27	135	6.7	1268	US-09-070-927A-591	Sequence 591, App
28	130	6.5	800	US-09-974-300-2879	Sequence 2879, Ap
29	129.5	6.4	9025608	US-10-156-761-1	Sequence 1, Appl1
30	124	6.2	966	US-10-156-761-3002	Sequence 3002, Ap
31	117.5	5.8	669	US-09-974-300-7149	Sequence 7149, Ap
32	116	5.8	3097	US-10-128-714-354	Sequence 354, App
33	116	5.8	3130	US-10-128-714-354	Sequence 5354, App
34	114.5	5.7	8911	US-10-044-090-17	Sequence 17, Appl1
35	113	5.6	888	US-09-815-242-3828	Sequence 3828, Ap
36	113	5.6	1089	US-09-815-242-6696	Sequence 6696, App
37	113	5.6	9212	US-09-070-927A-334	Sequence 334, App
38	112	5.6	885	US-09-815-242-9236	Sequence 9236, Ap
39	108.5	5.4	474	US-09-969-896-4	Sequence 4, Appl1
40	108	5.4	948	US-09-815-242-8764	Sequence 8764, Ap
41	108	5.4	13542	US-08-781-986A-154	Sequence 154, App
42	107.5	5.3	1338	US-10-102-806-177	Sequence 177, App
43	106.5	5.3	2837	US-09-822-846-58	Sequence 58, Appl1
44	105	5.2	378	US-09-770-791-215	Sequence 215, App
45	103.5	5.1	650	US-09-974-300-2884	Sequence 2884, Ap

ALIGNMENTS

RESULT 1
US-09-784-810A-1
; Sequence 1, Application US/09784810A
; Patent No. US20020082203A1
; GENERAL INFORMATION:
; APPLICANT: RASTRELLI, LUCIA
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 10716-08
; CURRENT APPLICATION NUMBER: US/09/784, 810A
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/182,360
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 60/191,261
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-09-784-810A-1
Alignment Scores:

QY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuThrAsnGlySerLeu 140
|||||
DB 361 AACCATTAATGCGGTATGAGAGAGTCAACCAATGAAGACCTCTTACCACTGACGCGTA 420
QY 141 LeuLeuGlySerProValLeuSerProMetLsnLeuSerLeuHisThrAlaSerGly 160
|||||
DB 421 TTGCGTGTCCCGCGGCTGCTGTCACCATGAACCTGCTGTCTCTGACACGCGCTCGGG 480
QY 161 LeuArgSerPheSerValLeuSerLeuLarTrpGlyPheIleAlaAspValAspLeuGlu 180
|||||
DB 481 CTGCGCCCTCTCTCTGTGCTCAGCCCTGGCGCTGGGCGCTTCATCTGTGATGTGACCTAAG 540
QY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
|||||
DB 541 AGTGAAGATATCGGCGCTGCGGAGATGCGCTTCACTCTGTGGACCTTCTGCGTCTG 600
QY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
|||||
DB 601 GCAGCCCTGCGCACCTACCGCGCGGCGGCGCTGCGCTACCTCCGTGAGAAAGTGGGTTCC 660
QY 221 LysThrProAlaSerProValValValGlnGlnGlnGlnProValAspAlaHisIleValPro 240
|||||
DB 661 AAGACACCTGCTCCCGCTGTGTGTGTCCAGAGGCGCGGAGATGATGACACCTGTGTGCA 720
QY 241 LeuGluGluGlnValProSerHisTyrGlnValValProAspGluAspPheValLeuVal 260
|||||
DB 721 CTGGAGAGAGCCAGTCCCTCTCTACAGACAGTGTGCGCGGACGAGAGACTTGTGTAGTC 780
QY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgGly 280
|||||
DB 781 CTGGACACTGCTGCACCTCCACCTGGGAGATGTTGTGTCGACCATGGGCGGCTGT 840
QY 281 AlaAlaGlyValMetHisLeuPheThrValArgAlaGlyValSerArgAlaMetLeu 300
|||||
DB 841 GCAGCTGGCGTATCATCTGTCTACAGTGGCGGAGGAGTGTCTGCTGCTGCTGCTG 900
QY 301 ArgLeuPheLeuAlaMetGlyGlyArgHisMetGlyTyrGlyGlyCysProTyrLeuVal 320
|||||
DB 901 CGCCCTTCTGCTGGCATGAGAGAGGCGAGCATATGAGATATGATGCCCTTGTGTA 960
QY 321 TyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal 340
|||||
DB 961 TATGTGCGCGTGTGCTGCTGCTGCTGAGCCCAAGATGAGGAAAGTGTGTTGTCAGTG 1020
QY 341 AspGlyGluMetValSerGluAlaValGlnGlnGlnValHisProAspThrPheTrp 360
|||||
DB 1021 GATGGGAAATGATGATTAGCAGGCGCGTGCAGGCGCGAGTGCACCCAAACTACTTCTGG 1080
QY 361 MetValSerGlyCysValGluProProPheSerTyrLysProGlnGlnMetProProPro 380
|||||
DB 1081 ATGTCACGCGTGTGCTGAGAGCGCGCGGCGGAGCTGGAAGCCCGACAGATGCGACCGCA 1140
QY 381 GlnGluProLeu 384
|||||
DB 1141 GAAGAGCCCTTA 1152

RESULT 3
US-10-023-282-90
Sequence 90, Application US/10023282
Publication No. US20030092893A1
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/10/023,282
CURRENT FILING DATE: 2001-12-20
EARLIER APPLICATION NUMBER: 09/205,258
EARLIER FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15

```

: EARLIER APPLICATION NUMBER: 60/094,657
: EARLIER FILING DATE: 1998-07-30
: NUMBER OF SEQ ID NOS: 1227
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 90
: LENGTH: 1533
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (112)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (123)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1522)
: OTHER INFORMATION: n equals a,t,g, or c
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1527)
: OTHER INFORMATION: n equals a,t,g, or c
: US-10-023-282-90

Alignment Scores:
Pred. No.: 3,13e-235 Length: 1533
Score: 1927.00 Matches: 371
Percent Similarity: 96.88% Conservative: 2
Best Local Similarity: 96.36% Mismatches: 11
Query Match: 9 Gaps: 0

US-09-784-810A-2 (1-384) x US-10-023-282-90 (1-1533)
QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
Db 73 ATGGATCCAGCGGGGCGGGGCGGGGCGGCTGCTCCCGCGGGCTGCGGTGCTGCTG 132
QY 21 LeuAsnProArgGlyGlyGlyValLeuGlnLeuPheArgSerHisValGlnPro 40
Db 133 CTGAACCGCGCGCGCGCAAGGCGCTTGCAGCTTCCGAGTCCAGTCCAGGCC 192
QY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuTrpGluArgAsnHisAla 60
Db 193 CTTTGGCTGAGGCTGAAATCTCTTCCAGCTGATGCTCAGCTGAGCGGCGGAGACCGC 252
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
Db 253 CGGGARCTGGTGGCGGAGAGAGCTGGGCGCTGGRACGCTCTGTGTCATGTYTGA 312
QY 81 AspGlyLeuMetHisGluValValAsnGlyLeu-MetGluArgProAspTrpGluTrpAl 100
Db 313 GACGGGCTGATGACGAGGTGGTGAACGGGCTTGAAGCGGCTGAGTGGAGACGC 372
QY 100 aIleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLe 120
Db 373 CATCCGAAAGCCCTGTGTAGCTCCAGCAGGCTCTGGCAACGSGGCGAGCTTCC 432
QY 120 uAsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuThrAsnGlyThrLe 140
Db 433 RAACCATTTATGGCTATRAGCAGGTACCAATGAACGCTCTGACCACTGACGCT 492
QY 140 uLeuLeuCysArgProValLeuSerProMetAsnLeuSerLeuHisThrAlaSerG 160
Db 493 ATGGCTGTGCGCGGCTGCTGTACACCAAGAACTGCTGTGCAACAGGCTTGGG 552
QY 160 yLeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeu 180
Db 553 GCTGCCCTCTTCTCTGCTGCTAGCTGCGCTGGGCTTCAATTCGATGAGTGAAGCT 612
QY 180 uSerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArg 200

```

```

Db 613 GAGTAGAAGATAGCGGCTGCTGGGAGATCGCTTCACTGTGGCACTTCCGGCT 672
QY 200 uAlaAlaLeuArgTrpTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPh 220
Db 673 GGCAGCCCTGGCAGCTACCGGCGGCGAGCTGCTAGCTAGTGAAGTGGGTTC 732
QY 220 eLysThrProAlaSerProValValGlnGlnGlyProValAspAlaHisLeuValPr 240
Db 733 CAAGACACCTGCTCCCGCTGTGTGTCCAGAGGCGCGGTAGTGAACACTGTGTC 792
QY 240 oLeuGluGlnGlnValProSerHisTrpGlnValValProAspGluAspPheValLeuVa 260
Db 793 ACTGAGAGGAGCAGGCGCTCTCACTGACAGATGTGTGCCAGCAGGACTTGTGCTAGT 852
QY 260 lLeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaIleProMetGlyArgCy 280
Db 853 CCTGGCACTGCTGCACTGCACTGGCGGAGTAGATGTTGTGTCACCAATGGGCGCTG 912
QY 280 sAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLe 300
Db 913 TGCAGCTGGCGTATCATCTGTCTACGTGGCGGAGTGTCTGTCATGCTGCT 972
QY 300 uArgLeuPheLeuAlaMetGluGlyArgHisMetGluTyrGluCysProTyrLeuVa 320
Db 973 GCGCCTTCTCTGCGCATGAGAGAGGCGGCAATGAGTGTGAATGCCCTACTTGT 1032
QY 320 lTyrValProValValAlaPheArgPheGluProLysAspGlyLysGlyValAlaPheAlaVa 340
Db 1033 ATATGTCGCCGTGTGTGCTTCCGCTTGGAGCCCAAGATGGGAAGGTGTGTCAGT 1092
QY 340 lAspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTr 360
Db 1093 GGATGGGAATGATGATTAGCAGAGCGCTGCAGGCGCAGTGTCCACCACTACTTCTG 1152
QY 360 pMetValSerGlyCysValGluProProProSerTrpLysProGlnMetProProPr 380
Db 1153 GATGGTCAGGGTGGTGGAGGCCCGCCAGCTGGAAGCCCAACAGATGCCAGCGC 1212
QY 380 oGluGluProLeu 384
Db 1213 AGAAGAGCCCTTA 1225

RESULT 4
US-09-784-810A-3
: Sequence 3, Application US/09784810A
: Patent No. US20020082203A1
: GENERAL INFORMATION:
: APPLICANT: RASTELLI, LUCA
: TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
: FILE REFERENCE: 10716-08
: CURRENT APPLICATION NUMBER: US/09/784,810A
: PRIOR FILING DATE: 2001-02-14
: PRIOR APPLICATION NUMBER: 60/182,360
: PRIOR FILING DATE: 2000-02-14
: PRIOR APPLICATION NUMBER: 60/191,261
: PRIOR FILING DATE: 2000-03-22
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3
: LENGTH: 1759
: TYPE: DNA
: ORGANISM: Mus musculus
: US-09-784-810A-3

Alignment Scores:
Pred. No.: 8.97e-193 Length: 1759
Score: 1594.50 Matches: 303
Percent Similarity: 88.25% Conservative: 35
Best Local Similarity: 79.11% Mismatches: 44
Query Match: 79.13% Indels: 1
Db: 10 Gaps: 1

```


US-09-784-810A-2 (1-384) x US-09-784-810A-3 (1-1759)

```
Oy 1 MetaspProalaglygylProargGlyValLeuProarProCysArsgValLeuValLeu 20
Db 327 ATGGAACCCAGTAGAATGCCCTCGAGAGACTGCTCCACAGGCCATGCAAGAGTGTCTGTG 386
Oy 21 LeuAnProarArgglygylValLeuValLeuValLeuValLeuValLeuValLeuValLeu 40
Db 387 CTGAACCCAGAGGTGGCAGAGGCTGCTGAGCTCTGCAAGCTCTGCAAGCCGTGTCAGGCC 446
Oy 41 LeuValaIagluValaIuIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
Db 447 TTCCTGAGAGGAGCAGATGAATCACTTAAGTACTGATACACCGAAGCAAGCAAGCC 506
Oy 61 ArgGluLeuValaIarSerGluGluLeuGlyArgTrpAspAlaValaIaMetSerGly 80
Db 507 AGGAGGCTGTGTGTGAGAGAGCTGGTCACTGGAGCCCTGGCAGTCAATGTCCGCT 566
Oy 81 AspGlyLeuMetHisGluValaIaAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
Db 567 GATGCTGTGATGATGAAGTGGTAATGGTAATGCAAGCCGAGAGCTGGAGAGCTGCC 626
Oy 101 IleGluIysProLeuCysSerLeuProAlaIaIysSerGlyAsnAlaIaLeuAlaSerLeu 120
Db 627 ATCCAGAAACCCCTGTGAGCCCTCCGAGAGCTCCGCAATGGCGTGGCAGCTTGTGTG 686
Oy 121 AsnHisTrpAlaGlyTrpGluGluValaIaThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
Db 687 AACCATATGCTGGGTGAGCAGACAGTGAATGATGACCTGCTCATCAATGCACTGACACTG 746
Oy 141 LeuLeuCysArsgProValaIaLeuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160
Db 747 CTGTTGTGCGCCCGCCGCTGTCCACATGAACCTGCTGCTCCGCAACCTGCTTGTGG 806
Oy 161 LeuArgSerPheSerValaIaLeuSerLeuAlaIaTrpIlePheIaIaAspValaIaSerGlu 180
Db 807 CTGGCGCTGTATCTGTGCTCAGTCTGCTGGGCGTTGTGCTGAGCTGGAGCTGGAG 866
Oy 181 SerAspIysTrpAlaGlyGluGlyLeuMetArgPheThrLeuGlyThrPheLeuArgLeu 200
Db 867 AGTGAAGATACAGAGCCCTGGGAGATTCGTTTCAACAGTGGCACCCTTGTCCGCTA 926
Oy 201 AlaAlaLeuArgTrpThrTrpArgGlyArgLeuAlaIaThrLeuProValaIaArgValaIaGlyPhe 220
Db 927 GCAAGCTGCCCATCTACACAGGCCCACTGGCTACCTCTGTAAGAACCTGTGGCTGT 986
Oy 221 LysThrProAlaSerProValaIaValaIaGluGluGlyProValaIaAspAlaHisLeuValaIaPro 240
Db 987 AAGAGACCCGCTCTCT--ACACTGTGTGAGAAAGGCGCCGTCGACACACACTTGTCTCT 1043
Oy 241 LeuGluGluGluValaIaProSerHisTrpGluValaIaIaProAspGluAspPheValaIaValaIa 260
Db 1044 CTGAGAGGAGCAGAGCTCTTCATGTGACTGTGTGATGCTGACAGAACAGAACTTGTCTGTG 1103
Oy 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaIaIaProMetGlyArgCys 280
Db 1104 CTGTGTGCTGTACACACACCACTGAGCTCCGAGCTGTTTGACAGACCACTATGGCGCTGT 1163
Oy 281 AlaAlaGlyValaIaMetHisLeuPheTrpValaIaArgAlaGlyValaIaSerArgAlaMetLeuLeu 300
Db 1164 GAGCTGTGTGTATGCAATCTGTCTAGCTAGCTGCGGGGTGTCAAGGCTGCGCTGTG 1223
Oy 301 ArgLeuPheLeuAlaMetGluGlyArgHisMetGluTrpGluCysProTrpLeuValaIa 320
Db 1224 CGCCTCTTCTGGCCATGCAAGAGGCAAGCATATGGAATCTGTACTGTCCATCTGCTGTT 1283
Oy 321 TyrValProValaIaIaPheArgLeuGluProIysAspGlyLysGlyValaIaPheAlaValaIa 340
Db 1284 CATGTGCGCGGTGTGCTTCCGCTGAGCCCAAGAGCCAGAGGCGGTGTCTTCTGTG 1343
Oy 341 AspGlyLeuLeuMetValaIaSerGluAlaValaIaGluGlyValaIaIaProAsnTrpPheTrp 360
```

```
Db 1344 GATGAGAGCTGATGATGAAGCTGTGCAAGGCCCAAGTGCACCACTACCTTGTG 1403
Oy 361 MetValaIaSerGlyCysValaIaGluProProSerTrpIysProGluGluMetProPro 380
Db 1404 ATGTGCTGTGGCAGACAGAGATGCCATTCGCCCGGAGACTCCGCGGGGCGCCACTCA 1463
Oy 381 GluGluPro 383
Db 1464 GAAGAACCA 1472

RESULT 5
US-09-970-516-5
: Sequence 5, Application US/09970516
: Patent NO. US20020099029A1
: GENERAL INFORMATION:
: APPLICANT: NO. US20020099029A1attis AG
: TITLE OF INVENTION: Induction of blood vessel formation through administration of
: FILE REFERENCE: 4-31617
: CURRENT APPLICATION NUMBER: US/09/970, 516
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 5
: LENGTH: 1149
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(1149)
: OTHER INFORMATION:
US-09-970-516-5

Alignment Scores:
Pred. No.: 3,67e-192 Length: 1149
Score: 1587.50 Matches: 302
Percent Similarity: 87.99% Conservative: 35
Best Local Similarity: 78.85% Mismatches: 45
Query Match: 78.78% Indels: 1
Db: 10 Gaps: 1

US-09-784-810A-2 (1-384) x US-09-970-516-5 (1-1149)
Oy 1 MetaspProalaglygylProargGlyValLeuProarProCysArsgValLeuValLeu 20
Db 1 ATGGAACCCAGTAGAATGCCCTCGAGAGACTGCTCCACAGGCCATGCAAGTGTCTGTG 60
Oy 21 LeuAnProarArgglygylValLeuValLeuValLeuValLeuValLeuValLeuValLeu 40
Db 61 CTGAACCCAGAGGTGGCAGAGGCTGCTGAGCTCTGCAAGCTCTGCAAGCCGTGTCAGGCC 120
Oy 41 LeuValaIagluValaIuIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
Db 121 TTCCTGAGAGGAGCAGATGAATCACTTAAGTACTGATACACCGAAGCAAGCAAGCCATG 180
Oy 61 ArgGluLeuValaIarSerGluGluLeuGlyArgTrpAspAlaValaIaMetSerGly 80
Db 181 AGGAGAGCTGTGTGTGAGAGAGTGGTCACTGGAGCGCCCTGGCAGTCAATGTCCGCT 240
Oy 81 AspGlyLeuMetHisGluValaIaAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
Db 241 GATGCTGTGATGATGAAGTGGTAATGGCTTAATGCAAGCCGAGAGCTGGAGAGCTGCC 300
Oy 101 IleGluIysProLeuCysSerLeuProAlaIaIysSerGlyAsnAlaIaLeuAlaSerLeu 120
Db 301 ATCCAGAAACCCCTGTGAGCCCTCCGAGAGCTCCGCAATGGCGTGGCAGCTTGTGTG 360
Oy 121 AsnHisTrpAlaGlyTrpGluGluValaIaThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
Db 361 AACCATATGCTGGGTGAGCAGACAGTGAATGATGACCTGCTCATCAATGCACTGACACTG 420
Oy 141 LeuLeuCysArsgProValaIaLeuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160
```


Oy 215 ----- 215
 Db 1587 GGCCTCCCTGAGCCCTGCTGACCTGCTCCTCAATGTTGTGTCCAGAGCTGACTGGA 1646
 Oy 215 ----- 215
 Db 1647 GACTGGGAGAGAGCTGGGATGACACTCTGTCCAGACCCACTGCTGCTTCAATCCCC 1706
 Oy 216 ----- Gly 216
 Db 1707 AACGCTCTCAAAACAGCTCAGCTTTCACCCATGCTGAGGCCCCAGAAATCCAGCA 1766
 Oy 217 ArgValGlyPheLysThrPro-----AlaSerProValValGlnGlnGlyProVal 234
 Db 1767 TCTTCGGGGTTCCTGCTCCACCCAGAGTCCCAAGAGCTTACCTGGGGCCCACTG 1826
 Oy 235 AspaLahLsLeuValProLeuGlnGlnValProSerHisTrpGlnValValProAsp 254
 Db 1827 GACCACCTCTCCCTCCCTGGGCTCTCCACTGCCCCAAGACTGGTGACAAATAGAGGG 1886
 Oy 255 GlnAspPheValLeuValLeuAlaLeuHisSerHisLsLeuAlaSerGlnMetPheAla 274
 Db 1887 GAG--TTGTACTCATGTTGGGAGCTTTCAGAGCCACTGCGCAGACCTGATGGCA 1943
 Oy 275 AlaProMetGlyArgCysAlaAlaGlyValMetHisLeuPheTyValArgAlaGlyVal 294
 Db 1944 GCCCAGCATGACACCTTTGATGATGAGCGCTTGTGACACCTGTTGGTGGAGGCGCATC 2003
 Oy 295 SerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGlnLysGlyArgHisMetClnTyx 314
 Db 2004 TCAGGGGCTGCATCTTACGACATTTTCTGGCCATGAGCAGTGAACCACTTACAGCCTG 2063
 Oy 315 GlnCysProTyLeuValTyValProValValAlaPheArgLeuGlnProLysAspGly 334
 Db 2064 GCGTGGCCCCATCTGGGCTATGCTGACAGCAGTGCCTTCCGCTTGAACCTCAAGCCT 2123
 Oy 335 LysGlnValPheAlaValAspGlyGlnMetValSerGlnAlaValGlnGlyGlnVal 354
 Db 2124 CGTGCGCTGCTCAGTATGATGAGGAGTATGAGTATGAGGCAATACAGCGCAGGTG 2183
 Oy 355 HisProAsnTyrrPheTrpMetValSerGlyCysValGlnProProSerTrpLysPro 374
 Db 2184 CACCAAGCTCTGCCAGCGCTGCTCAGTGG-----CCTGAGGTCAAAAGCCA 2231
 Oy 375 Gln 375
 Db 2232 CAA 2234
 RESULT 7
 US-09-970-516-3
 : Sequence 3, Application US/09970516
 : Patent No. US20020099029A1
 : GENERAL INFORMATION:
 : APPLICANT: NO. US20020099029A1a1s AG
 : TITLE OF INVENTION: Induction of blood vessel formation through administration of
 : FILE REFERENCE: 4-31617
 : CURRENT APPLICATION NUMBER: US/09/970, 516
 : NUMBER OF SEQ ID NOS: 6
 : SOFTWARE: Patent version 3.1
 : SEQ ID NO 3
 : LENGTH: 1857
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (1)..(1857)
 : OTHER INFORMATION:
 : US-09-970-516-3
 Alignment Scores:

	Pred. No.:	1,06e-98	Length:	1857
	Score:	856.50	Matches:	189
	Percent Similarity:	51.43%	Conservative:	63
	Best Local Similarity:	38.57%	Mismatches:	111
	Query Match:	42.51%	Indels:	127
	DB:	10	Gaps:	6

US-09-784-810a-2 (1-384) x US-09-970-516-3 (1-1857)
 Oy 10 ValLeuProAlaArgProCysArgValLeuValLeuAsnProAlaArgGlyLysGlyLys 29
 Db 418 CTGCTACCTCGCGCCGCCCGCTTCTTATGTGTAATCCCTTGGGGGCGGCGCTG 477
 Oy 30 AlaLeuGlnLeuPheArgSerHisValGlnProLeuAlaGlnAlaGlnLsSerPhe 49
 Db 478 GCGTCGACAGTGTATAGAACCACTGCTTCCATGATCTGTGAAGCTGGCTCTCC 537
 Oy 50 ThrLeuMetLeuThrGlnArgArgAsnHisAlaArgGlnLeuValArgSerGlnGlnLeu 69
 Db 538 AACCTCATCCAGACAGAAAGCAGACACACCGCGGAGTGTCCAGAGGGCTGAGCCTG 597
 Oy 70 GlyArgTrpAspAlaLeuValValMetSerGlyLysGlyLeuMetHisGlnValAlaAsn 89
 Db 598 AGTAGTGGAGTGGATGCAATGCTCAGCGTCCGGAGACGGGCTGTCAATGAGTCTGAAAC 657
 Oy 90 GlnLeuMetGlnArgProAspTrpGlnThrAlaLeuGlnLysProLeuCysSerLeuPro 109
 Db 658 GGGCTCTCAATGCCCTCTGACTGAGGAGAACCTGTGAAGATGCTGTGGCATCTTCCC 717
 Oy 110 AlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisTyAlaGlyTyGlnGlnVal 129
 Db 718 TGCGGCTCGGCGCAACGGCGCTGGCGGACAGTGAACCAAGCGGGGATTTGAGCCAGCC 777
 Oy 130 ThrAsnGlnAspLeuLeuThrAsnCysThrLeuLeuLeuCysArgProValLeuSerPro 149
 Db 778 CTGGGCTCGACCTGTGCTCACTGCTCAGCTGTGCTGTGCTGGGGTGTGGCCACC 837
 Oy 150 MetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSerLeu 169
 Db 838 CTGACACGCTCTCCGTAAGCCTGAGCGCTCGGGCTCCCGCTTCTCTCTCTCTCTCTCT 897
 Oy 170 AlaTrpGlyPheLeuAlaAspValAspLeuGlnSerAspLysTyArgArgLeuGlyGln 189
 Db 898 GCGTGGGGCTTCGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 957
 Oy 190 MetArgPheThrLeuGlyThrPheLeuArgLeuAlaLeuArgTyThrTyArgGlyArg 209
 Db 958 GCCGCTTCACACTGGGACGCGTGGCTCGCCACACTGACACTACCGCGGAGCGC 1017
 Oy 210 LeuAlaThrLeuProVal----- 215
 Db 1018 CTCTCTCACTCCCGCCGACCTGTGGAACCTGCTGCCACCTCGCCCATAGCTGTGCT 1077
 Oy 215 ----- 215
 Db 1078 CGTGCAAGTGGAGCTGACCTTAACCCAGACAGCCGCGCCCATGAGCCCACTACACC 1137
 Oy 215 ----- 215
 Db 1138 CTGCATGTTCTGTGTGCTGACCTGCTTCTCCCTGCCCCAGCTGGCCTGTCTCT 1197
 Oy 215 ----- 215
 Db 1198 GCGTCGCCAGAACCCCTGCCCCATCTGTCTCTCAACGTTGGGGGCCAGAGCTGGCTGG 1257
 Oy 215 ----- 215
 Db 1258 GACTGGGGTGGGGGATGCTCCGCTGCCCCGAGCCACTGTGTCTTACACTCTCT 1317
 Oy 216 -----GlyArgValGlyPheLysThrProAlaSer-----ProValValValGln 230
 Db 1318 GGCCTCTCCCAAGGAGCTCTCACTCAACCCGCTTCGCAAGGGGCGCCCGTAAATTCGCCCA 1377

QY 231 Gln-----GlyPro 233
Db 1378 TCCTGGGGTCCCACTTCCACCCCTGATGCCGGGTAGGGGCTCCACCTGCGGCCG 1437
QY 234 ValAspAlaHisLeuValProLeuGluGluGlnValProSerHisTrpGlnValPro 253
Db 1438 CCCGACCACTGGTCCCTCCGCTGGGACCCCGCCCTCCACAGACTGG--GTGACGCTG 1494
QY 254 AspGluAspPheValLeuValLeuAlaLeuLeuHisSerHisLeuHisLeuAspGluMetPhe 273
Db 1495 GAGGGGAGCTTTGTGCTCATGTTGGCCATCTGCGCCAGCCAGCTAGGCGCTGACCTGGTG 1554
QY 274 AlaAlaProMetGlyArgCysAlaAlaGlyValMetHisLeuPheTrpValAlaGlnGly 293
Db 1555 GCACCTCCGATGGCGCTTCGACAGAGCGCTGTGTGACCTGTGGGGGTGCTACCGGC 1614
QY 294 ValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMetGluLysGlyArgHisMetGlu 313
Db 1615 ATCTGGGGGCTGGCTGCTGCTGCGCTTTCTTGCCCATGAGCGTGTACCACTTCAGC 1674
QY 314 TyrGluCysProTyrLeuValTyrValProValAlaAlaPheArgLeuGluProLysAsp 333
Db 1675 CTGGGCTGTCGCGACGTGGGTACGCGCGGCCGCTTCCTCCGCTAGAGCGCTCACA 1734
QY 334 GlyLysGlyValPheAlaValAspGlyGluLeuMetValSerGluAlaValGlnGlyGln 353
Db 1735 CCACGCGGCTGCTCAGACGTGGAGCGGAGAGGTGAGATGAGGCGCGCTACAGGCACAG 1794
QY 354 ValHisProAsnTyrPheTrpMetValSerGlyCysValGluProProSerTrpLys 373
Db 1795 ATGACCTCGGCATCGGTACACTGCTCTGCTGGG-----CCTCGTGGCTGC--- 1839
QY 374 ProGlnGlnMetProProProGluGluPro 383
Db 1840 -----CCGGGGCGGGAGGCC 1854
RESULT 8
US-09-817-676A-13
Sequence 13, Application US/09817676A
Patent No. US20020042101A1
GENERAL INFORMATION:
APPLICANT: Spiegel, Sarah
APPLICANT: Kohama, Takafumi
TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
FILE REFERENCE: 00170/HG
CURRENT APPLICATION NUMBER: US/09/817,676A
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/194,318
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 2380
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (7)..(1860)
PUBLICATION INFORMATION:
TITLE: Molecular cloning and functional characterization of a
JOURNAL: J. Biol. Chem.
VOLUME: 275
ISSUE: 26
PAGES: 19513-19520
DATE: 2000-06-30
DATABASE ACCESSION NUMBER: AF245447
DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-13
Alignment Scores: 1.54e-98 Length: 2380
Pred. No.:

Score: 856.50 Matches: 189
Percent Similarity: 51.43% Conservative: 63
Best Local Similarity: 38.57% Mismatches: 111
Query Match: 42.51% Indels: 127
DB: 10 Gaps: 6
US-09-784-810A-2 (1-384) x US-09-817-676A-13 (1-2380)
QY 10 ValLeuProAlaGlyProCysArgValLeuValLeuLeuAsnProArgGlyGlyLysGlyLys 29
Db 424 CTGTACCTCGGCGCCCGCCGGTGTCTGTATTGTCATTCCTTTGGGGGTGGGGGCTG 483
QY 30 AlaLeuGlnLeuPheArgSerHisValGlnProLeuLeuAlaGluAlaGluLeuSerPhe 49
Db 484 GCGTGGCAGTGGTGTAAACACAGCTGCTCCCATGATCTGTGAAGTGGGCTGCTCC 543
QY 50 ThrLeuMetLeuThrGluArgArgAsnHisAlaArgGluLeuValArgSerGluLeu 69
Db 544 AACCTATCCAGAACACAGAACAGAACACCGCGGAGGTGTCCAGGGGCTAGCCTG 603
QY 70 GlyArgTrpAspAlaLeuValValMetSerGlyAspGlyLeuMetHisGluValAlaAsn 89
Db 604 AGTAGTGGAGATGCGATGCTACAGGCTCTCGGAGAGCGGCTGCTCATGAGGTGCTGAC 663
QY 90 GlyLeuMetGluArgProAspTrpGluThrAlaIleGlnLysProLeuCysSerLeuPro 109
Db 664 GGGCTCTGATGCGCCCTGACTGAGGAGAGTGTGAAGATGCTGTGGGCTGCTCC 723
QY 110 AlaGlySerGlyAsnAlaLeuAlaLeuAlaSerLeuAsnHisTyrAlaGlyTyrGluGlnVal 129
Db 724 TGGGCTCGGGACAGCGCTGGCCGAGAGAGTGAACAGACAGGGGGATTGAGGCACCC 783
QY 130 ThrAsnGluAspLeuLeuThrAsnCysTrpHisLeuLeuCysArgProValLeuSerPro 149
Db 784 CTGGGCTCGACCTGTGCTCAACTGCTACTGCTGTGCGGGGTGGGCCACCA 843
QY 150 MetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSerLeu 169
Db 844 CTGGACCTGCTCCCTGACCGCTGCGGTGGGTCCCGCTTTCTCTCTCTGCTG 903
QY 170 AlaTrpGlyPheIleAlaAspValAspLeuGluSerAspLysTyrArgArgLeuGlu 189
Db 904 GCGTGGGCTGTGTGTCAGATGTGATATCCAGAGGAGCGCTGAGGGCTGGGCACT 963
QY 190 MetArgPheThrLeuGlyThrPheLeuAlaGluAlaLeuArgThrTyrArgLysArg 209
Db 964 GCCCGCTTCACACTGGGACAGGTGGGCTGCGCACACTGACACTACCGCGGAGCG 1023
QY 210 LeuAlaThrLeuProVal----- 215
Db 1024 CTCTCTACTCCCGCGCACTGTGGAACCTGGCTGCCACCCCTGCCATAGCCTGCT 1083
QY 215 ----- 215
Db 1084 CGTGCAGATCGAGCTGACCTTAACCCAGACCCCGCCCATGGCCCACTACCC 1143
QY 215 ----- 215
Db 1144 CTGCATGCTGTGTGTCTGACTGCTCTTCCCTGCCAGCTGCTGCTGCTGCT 1203
QY 215 ----- 215
Db 1204 GCGTCGCAAGAACCCCTGCCCATCCTGTCCCTCAAGGTGGGGCCAGAGTGGCTGGG 1263
QY 215 ----- 215
Db 1264 GACTGGGTGGGGCTGGGAGATGCTCCGCTGCGCCGAGCCAGCTGCTGCTTACTCTCT 1323
QY 216 -----GlyArgValGlyPheLysThrProAlaSer-----ProValValValGln 230
Db 1324 GCGTCTCCCAAGGAGCTACTACCTACCCGCTGCGAAGGGGCGCCGTAATTCGCCCA 1383
QY 231 Gln-----GlyPro 233


```
Db 192 TATAGCTCCACAGAGCTCTGTGGCAACGNNCTGGCAGTTTCTTNAACCATTAATGCTGCT 251
Oy 126 TTTGlu-GlnVal, ThrAsnGlu-AspLeuLeuThrAsnGlySerHisLeuLeuLeuCyAsp 145
Db 252 TATGAGCAGGTACCATGATGAACTCTGTGACCAANTGCAGTNAATGNGTGCCGC 311
Oy 145 roValleu--SerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerph 164
Db 312 CGGTGTGTGTTNAACCATGAACTGNTGTCTTTCGCAANAGTTTNGGGGTGGCCNTCTT 371
Oy 164 eSerValleu-SerLeuAlaTyrGlyPheHisLeuAlaAspValaAspLeuGluSerAspLysT 184
Db 372 CTTTGTGTGTNANCNTGGCMTGGGGTTTNAATGTTGATTTTNGGACCCGAGAGTTAGAACT 431
Oy 184 YTAAGArgLeuGlyGluMetArgPheThrLeuGlyThrPhe 197
Db 432 ATCGGGGTTTGGGGAATATGTTTAATTTGGGAATTTTC 472
```

```
RESULT 10
US-10-015-219-658
; Sequence 658, Application US/10015219
; Publication No. US2003008299A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493C1
; CURRENT APPLICATION NUMBER: US/10/015, 219
; CURRENT FILING DATE: 2002-03-02
; NUMBER OF SEQ ID NOS: 1739
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO: 658
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 218..229, 235, 263, 265, 271, 286, 289
; OTHER INFORMATION: n = A,T,C or G
US-10-015-219-658
```

```
Alignment Scores:
Pred. No.: 2,79e-37 Length: 296
Score: 365.00 Matches: 68
Percent Similarity: 89.74% Conservative: 2
Best Local Similarity: 87.18% Mismatches: 4
Query Match: 18.11% Indels: 4
DB: 9 Gaps: 1
```

US-09-784-810A-2 (1-384) x US-10-015-219-658 (1-296)

```
Oy 307 GlnValGlyArgHisMetGluTyrGluGlyProTyrLeuValTyrValProValAla 326
Db 7 AACAAAGGCAAGCATATGAGATGATGATCCCTCTGTTATGTTGCGCGTGC 66
Oy 327 PheArgLeuGluProLysAspGlyLysGlyValPheAlaValaAspGlyGluLeuMetVal 346
Db 67 TTCGGCTTGAGCCCAAGATGAGGAAAGTGTGTTGACGTGATGGGAATGATGTT 126
Oy 347 SerGlnAlaValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysVal 366
Db 127 AGCGAGGCGGTGAGGCGCAGGTGTCACCAACTACTTGTGATGTGTCAGCGGTGCGTG 186
Oy 367 GluProProSerTrpLysProGln-----GlnMetProProPro 380
Db 187 GAGCCCCCGCCCAACCTGGCGCGCAACCAACNCTAAGGCAANTTCACACACC 240
```

```
RESULT 11
US-09-777-564-658
; Sequence 658, Application US/09777564
; Patient No. US20020022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
```

```
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777,564
; CURRENT FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 1730
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO: 658
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(296)
; OTHER INFORMATION: n = A,T,C or G
US-09-777-564-658
```

```
Alignment Scores:
Pred. No.: 2,79e-37 Length: 296
Score: 365.00 Matches: 68
Percent Similarity: 89.74% Conservative: 2
Best Local Similarity: 87.18% Mismatches: 4
Query Match: 18.11% Indels: 4
DB: 10 Gaps: 1
```

US-09-784-810A-2 (1-384) x US-09-777-564-658 (1-296)

```
Oy 307 GlnValGlyArgHisMetGluTyrGluGlyProTyrLeuValTyrValProValAla 326
Db 7 AACAAAGGCAAGCATATGAGATGATGATCCCTCTGTTATGTTGCGCGTGC 66
Oy 327 PheArgLeuGluProLysAspGlyLysGlyValPheAlaValaAspGlyGluLeuMetVal 346
Db 67 TTCGGCTTGAGCCCAAGATGAGGAAAGTGTGTTGACGTGATGGGAATGATGTT 126
Oy 347 SerGlnAlaValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysVal 366
Db 127 AGCGAGGCGGTGAGGCGCAGGTGTCACCAACTACTTGTGATGTGTCAGCGGTGCGTG 186
Oy 367 GluProProSerTrpLysProGln-----GlnMetProProPro 380
Db 187 GAGCCCCCGCCCAACCTGGCGCGCAACCAACNCTAAGGCAANTTCACACACC 240
```

```
RESULT 12
US-09-796-692-2905
; Sequence 2905, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
```

```

PRIORITY APPLICATION NUMBER: US 60/218,950
PRIORITY FILING DATE: 2000-07-14
PRIORITY APPLICATION NUMBER: US 60/222,903
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: US 60/223,416
PRIORITY FILING DATE: 2000-08-04
PRIORITY APPLICATION NUMBER: US 60/223,378
PRIORITY FILING DATE: 2000-08-07
PRIORITY APPLICATION NUMBER: US 09/796,692
PRIORITY FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2905
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapiens
US-10-040-862-2905

Alignment Scores:
Pred. No.: 6.63e-37 Length: 199
Score: 360.00 Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.87% Indels: 0
DB: 9 Gaps: 0

US-09-784-810A-2 (1-384) x US-10-040-862-2905 (1-199)
QY 306 MetGluLysGlyArgHisMetLunrYrGluCysProYrLreuValTyrrValProValVal 325
Db 3 ATGGAAGAGGCGGCGATATGAGATGAAATGCCCCCTACTGTGATATGCCCCGTGTC 62
QY 326 AlaPheArgLengLunrProLysAspGlyLysGlyValPheAlaValAspGlyGluLeuMet 345
Db 63 GCCTTCGCTTGAGAGCCCAAGATGGAAGGTGTGTTGCAGTGCATGGGAAATTCATG 122
QY 346 ValSerGluAlaValGlnGlyLysValHisProAsnYrPheTrrPmetValSerGlyCys 365
Db 123 GTTAGGAGGAGCGGTGCAGGCGCCAGGTGCACCCAACTACTTCTGGATGCTCAGTGTTC 182
QY 366 ValGluProProPro 370
Db 183 GTGAGACCCCGGCC 197

RESULT 14
US-09-969-896-9
: Sequence 9, Application US/09969896
: Publication No. US2003012553A1
: GENERAL INFORMATION:
: APPLICANT: Kossida, Sophia
: TITLE OF INVENTION: Regulation of human SpHingosine
: TITLE OF INVENTION: Kinase-Like Protein
: FILE REFERENCE: 004974.00594
: CURRENT APPLICATION NUMBER: US/09/969,896
: CURRENT FILING DATE: 2001-10-04
: PRIORITY APPLICATION NUMBER: US 60/238,005
: PRIORITY FILING DATE: 2000-10-06
: PRIORITY APPLICATION NUMBER: US 60/214,113
: PRIORITY FILING DATE: 2001-08-23
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 1614
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-969-896-9

Alignment Scores:
Pred. No.: 7.17e-24 Length: 1614
Score: 268.50 Matches: 109
Percent Similarity: 41.23% Conservative: 72
Best Local Similarity: 24.83% Mismatches: 170
Query Match: 13.33% Indels: 88

```

DB: 9 Gaps: 15
US-09-784-810A-2 (1-384) x US-09-969-896-9 (1-1614)
OY 2 ASPROAlaGlyPro-ArgGlyValLeuProArgProCysArgValLeuValLeu 21
DB 351 GACCTGCGAGAGATGCTGAGAACATGACGTGACGACCAAAACATTTACTGTTAT 410
OY 21 uasnProArgGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 41
DB 411 CAACCCGTTGGAGGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 470
OY 41 uLeuAlaGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeuVal 61
DB 471 GTTACCTTACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 530
OY 61 gGluLeuValArgSerGluValLeuValArgTrpAspAlaValLeuValMetSerGlyAs 81
DB 531 GGAAGACTCTGTATGAGATTACATGACAAATACGAGGATGCTGTGTGCGGAGAG 590
OY 81 pGlyLeuMetHisGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 101
DB 591 TGGATGTTGAGGAG 650
OY 101 eGlnLys-----ProLeuCysSerLeuPr 109
DB 651 CGACCAAGACACCCCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 710
OY 109 oAlaGlySerGlyAsnAlaLeuAlaLeuAlaSerLeuAsnHisTyrAlaGlyTyrGluGln 129
DB 711 CGAGGGGTCAACGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 746
OY 129 LThrAsnGluValLeuValLeuValLeuValLeuValLeuValLeuValLeuValLeu 149
DB 747 GGGACACAG 804
OY 149 oMetAsnLeuSerLeuHisThrAlaSerGlyLeuValLeuValLeuValLeuValLeu 169
DB 805 -ATGAGATGCTGT 863
OY 169 uAlaTrpGlyPheLeuAlaAspValLeuValLeuValLeuValLeuValLeuValLeu 189
DB 864 GGGCTACGCTTCTACGAGATCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923
OY 189 uMetArgPheThrLeuGlyThrPheLeuValLeuValLeuValLeuValLeuValLeu 209
DB 924 TGCAGATACGAGATTTTGAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 983
OY 209 gLeuAlaThrLeuPro-----ValGlyArgValGlyPheLysThrPro----- 223
DB 984 AGTGTCTTCTGT 1043
OY 224 -AlaSerProValValLeuGlnGlnGlyProValAspAlaHisLeuValProLeuGluG 243
DB 1044 GGCAGAGATGCTTGT 1088
OY 243 uGlnValPro-----SerHisTrpGlnVal 251
DB 1089 GGAAGCAAGAAAGCACTGTATGTTGGAAGCTGCGAGAGAGAGAGAGAGAGAGAGAG 1148
OY 251 LValProAspGlyAspPheValLeuValLeuValLeuValLeuValLeuValLeuVal 271
DB 1149 CGT 1196
OY 271 uMetPheAlaAlaProMetGlyArgCysAlaAla-----GlyValMetHis 286
DB 1197 TTTGTGCGAG 1256
OY 286 sLeuPheTyrValArgAlaGlyValLeuValLeuValLeuValLeuValLeuValLeu 306
DB 1257 CTCATGCTCATGCGGAGAA---TGCTCCAGGTCAATTTTCTGAG---TTTCTCATGAG 1310
OY 306 tGlnLysGlyArgHisMetGlyTyrGlyCysProTyrLeuValTyrValProValAla 326

DB 1311 GCACACCAACAG 1370
OY 326 aPheArg-----LeuGluProLysAsp-----GlyLys 335
DB 1371 ATTCCAGTTTACCTGCAAGACATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 1430
OY 335 sGlyValPheAla----- 339
DB 1431 GAACCGCTTGGGACACATTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1490
OY 340 -----ValAspGlyGluLeuMetValSerGluAlaValGlnGlyValLeuValLeu 356
DB 1491 CTCTGGAAGTGGAG 1550
OY 356 oAsnTyrPheTrpMetValSerGlyCysValGluProProGlySerTrpSerTrpPro 374
DB 1551 CCAGCTGTTGCTGT 1599
RESULT 15
US-09-969-896-16
; Sequence 16, Application US/09969896
; Publication No. US2003012533A1
GENERAL INFORMATION:
; APPLICANT: Kossida, Sophia
; TITLE OF INVENTION: Regulation of human Sphingosine
; FILE REFERENCE: 004974.00594
; CURRENT APPLICATION NUMBER: US/09/969,896
; PRIOR APPLICATION NUMBER: US 60/238,005
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/314,113
; PRIOR FILING DATE: 2001-08-23
; NUMBER OF SEQ. ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 16
; LENGTH: 4413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-896-16
Alignment Scores:
Pred. No.: 3,28e-23 Length: 4413
Score: 268.50 Matches: 109
Percent Similarity: 41.23% Conservative: 72
Best Local Similarity: 24.83% Mismatches: 170
Query Match: 13.33% Indels: 88
DB: 9 Gaps: 15
US-09-784-810A-2 (1-384) x US-09-969-896-16 (1-4413)
OY 2 ASPROAlaGlyPro-ArgGlyValLeuProArgProCysArgValLeuValLeu 21
DB 426 GACCTGCGAGAGATGCTGAGAACATGACGTGACGACCAAAACATTTACTGTTAT 485
OY 21 uasnProArgGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 41
DB 486 CAACCCGTTGGAGGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 545
OY 41 uLeuAlaGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeuVal 61
DB 546 GTTACCTTACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 605
OY 61 gGluLeuValArgSerGluValLeuValArgTrpAspAlaValLeuValMetSerGlyAs 81
DB 606 GGAAGACTCTGTATGAGATTACATGACAAATACGAGGATGCTGTGTGCGGAGAG 665
OY 81 pGlyLeuMetHisGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 101
DB 666 TGGATGTTGAGGAG 725
OY 101 eGlnLys-----ProLeuCysSerLeuPr 109

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:10:08 ; Search time 65 Seconds
(without alignments)
1811.752 Million cell updates/sec

Title: US-09-784-810A-2
Perfect score: 2015
Sequence: 1 MDPAAGPGRGVLPKRCRLVL.....CPEPPPSWKPOQMPPEEPL 384

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-O=/cgn2_1/USPTO/spool/US09784810/runat_11072003_103442_26578/app_query.fasta.1.583
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09784810.ecgn1.1.40.ernuat_11072003_103442_26578 -NCPD=6 -ICPD=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOS
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

Database :

- Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Dackfiles1.seq:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	127.5	6.3	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
2	126.5	6.3	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
3	113	5.6	900	4 US-08-961-527-332	Sequence 332, App
4	110	5.5	1003	4 US-09-134-001C-1937	Sequence 1064, Ap
5	105	5.2	975	4 US-09-412-250-3	Sequence 1937, Ap
6	102.5	5.1	5092	4 US-09-754-250-1	Sequence 3, Appl1
7	99.5	4.9	111282	4 US-08-367-841A-43	Sequence 43, Appl1
8	98	4.9	22481	5 PCT-US95-07201-43	Sequence 43, Appl1
9	98	4.9	22481	4 US-09-875-223-2	Sequence 2, Appl1
10	98	4.9	4403765	4 US-09-103-840A-2	Sequence 1, Appl1
11	98	4.9	4403765	4 US-09-103-840A-1	Sequence 1, Appl1
12	98	4.9	4411529	4 US-09-103-840A-2	Sequence 1, Appl1

13	96	4.8	23673	4 US-09-773-816-1	Sequence 1, Appl1
14	95	4.7	36519	3 US-08-923-137-2	Sequence 2, Appl1
15	94	4.7	1649	2 US-08-845-566-2	Sequence 2, Appl1
16	93	4.6	5360	3 US-08-444-818-53	Sequence 53, Appl1
17	93	4.6	6785	3 US-08-444-818-65	Sequence 65, Appl1
18	93	4.6	7310	3 US-08-444-818-74	Sequence 74, Appl1
19	93	4.6	8316	3 US-08-444-818-88	Sequence 88, Appl1
20	93	4.6	8987	3 US-08-444-818-137	Sequence 137, App
21	93	4.6	9185	3 US-08-444-818-132	Sequence 122, App
22	93	4.6	9185	3 US-08-444-818-123	Sequence 123, App
23	93	4.6	9379	3 US-08-444-818-176	Sequence 176, App
24	93	4.6	9379	4 US-09-388-874-1	Sequence 1, Appl1
25	93	4.6	9401	1 US-07-910-760-9	Sequence 1, Appl1
26	93	4.6	9401	1 US-08-440-519-9	Sequence 9, Appl1
27	93	4.6	9401	1 US-08-440-519-9	Sequence 9, Appl1
28	93	4.6	9401	4 US-08-823-895A-25	Sequence 25, Appl1
29	92	4.6	2201	4 US-08-823-895A-25	Sequence 17, Appl1
30	92	4.6	3666	6 5248670-3	Patent No. 5248670
31	92	4.6	9401	5 PCT-US91-02225-9	Sequence 9, Appl1
32	91	4.5	49272	1 US-08-614-770A-1	Sequence 1, Appl1
33	90.5	4.5	3546	4 US-08-872-757-3	Sequence 3, Appl1
34	90.5	4.5	4041	4 US-09-105-537-36	Sequence 36, Appl1
35	90.5	4.5	35100	2 US-08-770-379-17	Sequence 17, Appl1
36	90.5	4.5	35100	4 US-08-757-660A-17	Sequence 17, Appl1
37	90.5	4.5	35100	4 US-09-230-371A-17	Sequence 17, Appl1
38	90.5	4.5	36778	4 US-09-105-537-5	Sequence 5, Appl1
39	90	4.5	3027	2 US-08-680-326-23	Sequence 23, Appl1
40	90	4.5	19056	4 US-09-272-032-8	Sequence 8, Appl1
41	89.5	4.4	2040	2 US-08-031-538-10	Sequence 10, Appl1
42	89.5	4.4	2769	4 US-09-118-408-1	Sequence 1, Appl1
43	89.5	4.4	2769	4 US-09-506-855-1	Sequence 1, Appl1
44	89	4.4	1907	4 US-09-443-184-40	Sequence 40, Appl1
45	88.5	4.4	1611	2 US-08-551-211-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-103-840A-2
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:

Prod. No.: 24
Score: 127.50
Percent Similarity: 41.50%
Best Local Similarity: 27.89%
Query Match: 6.33%
DB: 4
Length: 4403765
Matches: 41
Conservative: 20
Mismatches: 65
Indels: 21
Gaps: 2

US-09-784-810A-2 (1-384) x US-09-103-840A-2 (1-4403765)

Db 530 AGCAAGAAATTCGCCACCGGGTAGTCA----- 501
QY 201 AAlaAlaLeuArGhTyrArGlyArGleuAlaThrLeuProValGlyArGlyPhe 220
Db 500 -----GACCTTATCACTTAT 486
QY 221 LysThrProAlaSerProValValGlnGlnGlyProValAlaHis----- 237
Db 485 GCGGGAACCATGATGAGAGTGAATAACCAACCAAGAGATCCAAACGACGATAG 426
QY 238 ---LeuValProLeuGlu-GluGlnValProSerHisTrpGlnValValProAspGluAs 256
Db 425 ACGGCAAAATCTTTGAGGAAAGCTTCTCGTCAC-----TTGGTCCCA----- 380
QY 256 pPheValLeuValLeuAlaLeuHisSerHisLeuAlaSerGluMetPheAlaAlaPr 276
Db 379 -----ATGCGCCCGCATGATGGCAATAAACGCTTATATGCGTCC 342
QY 276 oMetGlyArGlyAlaAlaGlyValMetHisLeuPheTyrValAlaGly----- 293
Db 341 ATTAGCGGATTTGGAGGATGTAATAATGATGATGATATATTCGCTTTCAATCCATT 282
QY 294 -ValSerArGlyAlaMetLeuLeuArGlyLeuPheLeu--AlaMetGluGlyArGlyHisMe 312
Db 281 GGAAGCTCCACAAATGCTTACCACTCTTTACCAACCGATATCAAGCAACTCCAAATCT 222
QY 312 Tglu--TyrGluCysProTyrLeuValTyrValProValAlaAlaPheArGlyLeuGluPr 331
Db 221 GGAACCTATTAAGGCGCAAAACCTGATCAT-----GA 189
QY 331 oLysAspGlyLysGlyValPheAlaValAspGlyGlu-LeuMetValSerGluAlaValG 351
Db 188 GCGGAGACAGAGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 129
QY 351 InglyGlnVal-----HisProAsnTyrPheTrpMetValSerGlyValG 367
Db 128 TGAGATCCGACCTACGCGGCTCCTCAAGATTTTG--CTACGAAACCTACACTTG 72
QY 367 LuProProPseTrp--LysProGlnGlnMetProProGlu 381
Db 71 AAAGTCCACCTCTTCTAATAAACCAATAAATCCGCTAACCGGAG 25
RESULT 5
US-09-134-001C-1937
; Sequence 1937, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1937
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1937

US-09-784-810A-2 (1-384) x US-09-134-001C-1937 (1-975)
QY 16 ArgValLeuValLeuLeuAsnProArGlyGlyLysGlyLysAlaLeuGlnLeuPheArG 35
Db 34 CGTCAGAAATTTAT 81
QY 36 SerHisValGlnProLeuLeuAlaGlnAlaGlnLeuSerPheThrLeuMetLeuThrGlu 55
Db 82 CGT-----GTATTACCAATGCA-----CTGATTAACCTGAG 114
QY 56 ArgArGlnHisAlaArGlyLeuValArGlySerGlnGlnLeuGly----- 70
Db 115 AAGCAGGTTATGAAACGAGTCAATATGCAATGCAATGCAATGCAATGCAATGCAATGCA 174
QY 71 -----ArgTrpAspAlaLeuValValMetSerGlyAspGlyLeu 83
Db 175 GCTGAAGAGCACTAGAAAGTGAATATGATTACTTACCTGACCTGAGAGTGCAGTCA 234
QY 84 MetHisGlnValValAlaGlnGlyLeuMetGluArGlyProAspTrpGluThrAlaGlnLys 103
Db 235 TTAAATAGGTGCTCAACGGAATCCGCAACACCAAT-----CGGCT 279
QY 104 ProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeuAsnHisTyr 123
Db 280 AAATTAGGTATATACCAATGGGACCGTTATGACTTGGAGAGCACTTCAAT----- 333
QY 124 AlaGlyTyrGlnGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeuLeuGly 143
Db 334 -----TTACCA 339
QY 144 ArgProValLeuSerProMetAsn--LeuLeuSerLeuHisThrAlaSerGlyLeuArG 162
Db 340 AGCATATATATGGGGCGATGATGATATATATATATATATATATATATATATATATATAT 390
QY 163 SerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeuGlnSerAsp 182
Db 391 -----GTAAATATGCAAAATGAAATAT 414
QY 183 LysTyrArGlyArGlnGlyGlnMetArGlyPheThrLeuGlnThrPheLeuArGlnAla 202
Db 415 CGTTAT-----TTCAATTAACCTACTACTCA 438
QY 203 LeuArGhTyrArGlyArGlnGlnLeuAlaThrLeuProValGlyArGlyPheLysThr 222
Db 439 -----GAGGGGAACCTAAC-----CAATATCTTAATGAACA 471
QY 223 ProAlaSerProValValAlaGlnGlnGlyProValAlaHisLeuValProLeuGln 242
Db 472 CCA--AGTAAGTTGAATCAATTTGAGACCGTTC--GCGTATTAATTAAGGATTC 525
QY 243 GlnGlnValProSer-----HisTrpGlnValVal 252
Db 526 GAAATGTTACTCAATAATGAACAGTATGATGATGATGATGATGATGATGATGATGATGAT 585
QY 253 ProAspGlnAspPheValLeuValLeuAlaLeuHisSerHisLeuAlaSerGlnMet 272
Db 586 CAAGGAAACCTTATATATCTTTAGCTTTAAGATTAATGATGATGATGATGATGATGATGAT 645
QY 273 PheAlaAlaProMetGlyArGlyGlyAlaAlaGlyAlaMetHisLeuPheTyrValArGly 292
Db 646 TTA--GTTCCAGATCCGAGAGCTTGACGAGCTTATTCACGTTAAATTTTGAAGAAA 702
QY 293 GlyValSerArGlyAlaMetLeuLeuArGlyLeuPheLeuAlaMetGluLysGlyArGlyHisMet 312
Db 703 ---GCAAAATCTGCGAATTTGGTCAATTTATGACACTACCTGAGTGCAGTGCAGTGCAG 759
QY 313 GlnTyrGlnCysProTyrLeuValTyrValProValAlaAlaPheArGlyLeuProLys 332
Db 760 AAACAT-----CTTAAGTCAATTTATGCTTAAGCGAAGCTTAAATTTTCAATCA-- 810
QY 333 AspGlyLysGlyValPheAlaValAspGlyGlnLeuMetValSerGlnAlaValGlnGly 352
Db 811 -----TTTACTGATATGCACTTAATGATGATGATGATGATGATGATGATGATGATGAT 852

[illegible]

```

1  TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
2  TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
3  NUMBER OF SEQUENCES: 43
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: Morgan & Pinnegan
6  STREET: 345 Park Avenue
7  CITY: New York
8  STATE: New York
9  COUNTRY: USA
10 ZIP: 10154
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: floppy disk
14 COMPUTER: IBM PC Compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: WORDPERFECT 5.1
17
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/367,841A
20 FILING DATE: 30-DEC-1994
21 CLASSIFICATION: 435
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: 08/257,963
24 FILING DATE: 07-JUN-1994
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: 07/952,796
27 FILING DATE: 24-SEP-1992
28 ATTORNEY/AGENT INFORMATION:
29 NAME: DOROTHY R. AUTH
30 REGISTRATION NUMBER: 36434
31 REFERENCE/DOCKET NUMBER: 20264126052
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: (212) 758-4800
34 TELEFAX: (212) 751-6849
35 INFORMATION FOR SEQ. ID NO.: 43:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 22481 base pairs
38 TYPE: Nucleic Acid
39 STRANDEDNESS: Double
40 TOPOLOGY: Unknown
41 MOLECULE TYPE: Genomic DNA
42 FEATURE:
43 NAME/KEY: Pl-147
44 LOCATION:
45 IDENTIFICATION METHOD:
46 OTHER INFORMATION: full length genomic
47 OTHER INFORMATION: sequence for PEDF plus flanking sequences.
48
49 US-08-367-841A-43
50
51 Alignment Scores:
52 Pred. No.: 13.1 Length: 22481
53 Score: 98.00 Matches: 80
54 Percent Similarity: 33.96% Conservative: 28
55 Best Local Similarity: 25.16% Mismatches: 82
56 Query Match: 4.86% Indels: 129
57 Gaps: 17
58
59 US-07-784-810A-2 (1-384) x US-08-367-841A-43 (1-22481)
60
61 139 ThirleuenuCysArgProValLeuSerPrometAnleuSerleuHisThraIa 158
62 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 2024# ACAgTCGCTGCTGTCGAGCGAGGTGAGTCAAGTGCAGAGCACTCGGCTCACCAAC 2030303
64
65 159 Ser-----GlyLeuArgSerPheSerValLeuSerleuAlaTrpGlyPheIleAla 175
66 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 20304 TCCGCCCGCGCTGGGTAAACGCAATTCCTGCTCCACGCTC----- 20345
68
69 176 AsyValAspLeuGlnSerAspLysTyrArgArgLeuGlyGlnMetArgPheThrIleuGly 195
70 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
71 20346 -----CTGAGTACGCTGGGATTAACAGCGCTGCGCTCA-----TGCGTGGCT 20387
72
73 196 ThrPheLeuArgLeuAlaIleuArgThrTyrArgGlyArgLeuAlaThrIleuProVal 215
74 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
75 20388 AATTATATACAGCGGCTTCTTCAT-----GTT 20417

```



```

QY 216 GYATGVALGlypThrProAlaser-----ProValValaIngIngly 232
DB 20418 GGTGAGGTGGTCTCAAACTCCAACTCAGGTATCGGCTGGCTGCTCCAAAC 20477
QY 233 ProValAspAlaHisLeuValProLeuIngInglyProSerHisTrpGln----- 250
DB 20478 ACAGGGGATTCAGCATGAGCCA-----CCAGCTTGCCCAATCGTTGGCATTTAAAG 20531
QY 251 -----ValValProAspGlnAspPheValLeuValLeuAlaLeu----- 263
DB 20532 GCTTTCAGTGTACT-----GACTTCTT-TTAGTCTTAAGCTGTAACTGTAACTTT 20584
QY 264 -----Leu-HisSerHisLeuAlaSerGlnUmetPheAlaAlaPromet----- 277
DB 20585 CTTGGGCCAGCGCTATCACAGGATCTCTCGGGAATGTGACAGACAGTCCCAACCCG 20644
QY 278 -----GlyArgCysAlaAlaGlyValMetHisLeuPheTy----- 289
DB 20645 AGGAGACACCCAGGTGTGCA-----CACAGTTTCTGTCAACGATTTCGGA 20692
QY 290 -----ValArgAlaGlyVal 294
DB 20693 GGACTCTTGGAATCCCTGGAACACCATCTGTCCATGGAGCTTAGGTTAAGAGCCTCTGT 20752
QY 294 LserArgAlaMetLeuLeuArgLeuPheLeuAlaMetGlnUlysGlyArgHisMetGlnUty 314
DB 20753 TCAAGAGAGGCTTTTGTCTGTGTG-----GGTGGATGGGGGTAAAGT 20794
QY 314 rglu-----CysProTyLeuVal 320
DB 20795 CTCGAAGCCCTCTTRCGGSCCTTCGTATTCCTATNCCCGGTTCTGCCCTGTCTTGT 20854
QY 320 lTyValProVal----- 324
DB 20855 CCAGTCTCTCTATTAAACAATGAGCAGTAATGTACACCGATGGACTTTGGAGACAA 20914
QY 325 -----ValAlaPheArg-LeuGlnPro-----LysAspGlyUlysGlyValPheAlaVala 341
DB 20915 TAAAGACCTGATTTTCATTTCTAGCTCTTAAACACAGGACAGACATTTCTTACAGAAC 20974
QY 341 spgGlyLeuMetValSerGlnAlaValGlnGlyGlnValHisProAsnTyPheTrpM 361
DB 20975 AACTTCAGTGTGTATTAGCCAGTAGAAGGACACACATCTTTCTG----- 21027
QY 361 etValSerGlyCysValGlnProProSerTrpLysProGlnGlnMet 377
DB 21028 -----AAGAAACCTCAGAGATG 21045

```

RESULT 9 PCT-US95-07201-43

```

Sequence 43, Application PC/TUS9507201
GENERAL INFORMATION:
APPLICANT: Chader, Gerald J.; Becerra, Sofia
APPLICANT: Patricia; Schwartz, Joan P.;
APPLICANT: Tanikawa, Takayuki
TITLE OF INVENTION: PIGMENT EPITHELIUM
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07201

```

```

FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/367,841
FILING DATE: 30-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,963
FILING DATE: 07-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/952,796
FILING DATE: 24-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36434
REFERENCE/DOCKET NUMBER: 20264126PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 751-6849
TELEFAX: (212) 758-4800
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 22481 Base Pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: Pl-147
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: full length genomic
OTHER INFORMATION: sequence for PDF plus flanking sequences.
PCT-US95-07201-43
Alignment Scores:
Pred. No.: 13.1 Length: 22481
Score: 98.00 Matches: 80
Percent Similarity: 33.96% Conservative: 28
Best Local Similarity: 25.16% Mismatches: 82
Query Match: 4.86% Indels: 129
DB: 5 Gaps: 17
US-09-784-810A-2 (1-384) x PCT-US95-07201-43 (1-22481)
QY 139 ThrLeuLeuLeuGlyArgProValLeuSerProMetAlaLeuLeuSerLeuHisThrAla 158
DB 20244 ACAGTCTGCTGTGCGCCAGGCTGAGTCCAGTGCACGATCTCGCTCACCAACC 20303
QY 159 Ser-----GlyLeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAla 175
DB 20304 TCCGCCCCCGCTGGGTAAAGCATTCCTTCCCTCACGCTC----- 20345
QY 176 AspValAspLeuGlnSerAspLysTyArgArgLeuGlnGlyGlnMetArgPheThrLeuGly 195
DB 20346 -----CTGAGTAGCTGGGATTAACAGCGGTGCTCCA-----TGCTGGCT 20387
QY 196 ThrPheLeuAlaGlyLeuAlaLeuAlaLeuArgThrTyArgGlyArgLeuAlaThrLeuProVal 215
DB 20388 AATTTTATACAGACGCGGTTTCTCCAT-----GTT 20417
QY 216 GYATGVALGlypThrProAlaser-----ProValValaIngIngly 232
DB 20418 GGTGAGGTGGTCTCAAACTCCAACTCAGGTATCGGCTGGCTGCTCCAAAC 20477
QY 233 ProValAspAlaHisLeuValProLeuIngInglyProSerHisTrpGln----- 250
DB 20478 ACAGGGGATTCAGCATGAGCCA-----CCAGCTTGCCCAATCGTTGGCATTTAAAG 20531
QY 251 -----ValValProAspGlnAspPheValLeuValLeuAlaLeu----- 263
DB 20532 GCTTTCAGTGTACT-----GACTTCTT-TTAGTCTTAAGCTGTAACTGTAACTTT 20584
QY 264 -----Leu-HisSerHisLeuAlaSerGlnUmetPheAlaAlaPromet----- 277

```

Db 20585 CTTGGGCGACGGCTATCAACAGATCTCTGGGAATCTGACAGACAGTGCCTCAAAACCG 20644
 QY 278 -----GLYARGCYSALAALAGLYVALMETHISTLEUPHETYL----- 289
 Db 20645 AGGAGACACCGCCAGGTGTGCA-----CACAGCTTTCTGTCAAGATTTGGGA 20692
 QY 290 -----VALARGALAGLYVA 294
 Db 20693 GGACTCTGGGATCCCTGACACACATCTGTTCATGCGACCTTAGGTAAAGGCTCTGT 20752
 QY 294 1SERARGALAMETLEUARGLEUPHELEUALAMETGLULYSGLYARGHISMETGLUTY 314
 Db 20753 TCAAGGAGGCTTTGCTCTGTGTG-----GGTGGATGGGGAGAGT 20794
 QY 314 RGLU-----CYSPROTYLLEUVA 320
 Db 20795 CTCGAAGCCCTTTRCGGSCCTTGCGATTCCTATMCCCCGGTCTGCGCTGTAGT 20854
 QY 320 1TYVALPROVAL----- 324
 Db 20855 CCAAGTCTCTCTTTTAAACAATGACAGTAATGTACACCGATGACCTTTGGAGACAA 20914
 QY 325 -----VALALAPHEALY-LEUGLUPRO--LYSASPGLYLYSGLYVALPHEALAYLA 341
 Db 20915 TAAAGACCTGATTAATTAATCTTAAACCAACGAGAGAACATCTTTCACGACAC 20974
 QY 341 SPGLYGLUEUMETVALSERGLUALAVALINGLYGLUVALHISPROASNTYRPHETPRM 361
 Db 20975 AACTTCAGTTGATTATGAGCCAAAGGTAAAGAACGCAACGATCTCTTTCG----- 21027
 QY 361 ETVALSERGLYCYSVALGIUPROPROSERTRPLYSPROGLINGMET 377
 Db 21028 -----AAGAAACCTCAGGAGATG 21045
 RESULT 10
 US-09-875-223-2
 ; Sequence 2, Application US/09875223
 ; Patent No. 6391850
 ; GENERAL INFORMATION:
 ; APPLICANT: No. 6391850thwestern University
 ; APPLICANT: No. 63918501 Bouck
 ; APPLICANT: David Dawson
 ; APPLICANT: Paul Gillis
 ; TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis
 ; FILE REFERENCE: 0290-2303
 ; CURRENT APPLICATION NUMBER: US/09/875,223
 ; PRIORITY FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 09/122,079
 ; PRIOR FILING DATE: 1998-07-23
 ; PRIOR APPLICATION NUMBER: PCT/US98/15228
 ; PRIOR FILING DATE: 1998-07-23
 ; PRIOR APPLICATION NUMBER: US 08/899,304
 ; PRIOR FILING DATE: 1997-07-23
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 22484
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 1...22484
 ; OTHER INFORMATION: "n" means either a, c, t, or g
 US-09-875-223-2
 Alignment Scores:
 Pred. No.: 13.1 Length: 22484
 Score: 98.00 Matches: 80
 Percent Similarity: 33.96% Conservative: 28
 Best Local Similarity: 25.16% Mismatches: 129
 Query Match: 4.86% Indels: 17
 Db: 4 Gaps: 17

US-09-784-810A-2 (1-384) x US-09-875-223-2 (1-22484)
 QY 139 ThrLeuIleuLeuCysArgProValLeuSerProMetAspIleuLeuSerLeuHisThrAla 158
 Db 20244 ACAGTCTGCTCTGTCTGCGCAGGCGTGGAGTCCAGTGCACGATCTCGGCTCACCAAC 20303
 QY 159 Ser-----GLYLEUARGSERPHESEVALLEUSERLEUALATYRGLYPHEALALA 175
 Db 20304 TCGGCCCCCGTGGGTAAAGCATCTCTGCTCAAGCTC----- 20345
 QY 176 AspValAspLeuLeuSerAspLysTyrArgArgLeuGlyLeuMetArgPheThrLeuGly 195
 Db 20346 -----CTGAGTAGTGGGATTAACAGCGGTGCTSCA-----TGCCGTGCT 20387
 QY 196 ThrPheLeuArgLeuAlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProVal 215
 Db 20388 AATTATATACAGCAGCGGCTTCTCCAT-----GTT 20417
 QY 216 GLYARGVALGLYPHELEUHPHELEUHPROVALSER-----PROVALVALALINGLY 232
 Db 20418 GGTACGGCTGTCTCAAACTCCAACTCAGGTGATCGGCTGCTGCTCCCAAAAC 20477
 QY 233 ProValAspAlaHisLeuValProLeuGluGluValProSerHisTrpGln----- 250
 Db 20478 ACAGGGATTCAGGATGAGCA-----CCACGCTGCCAATCGTTGGCATTTGAG 20531
 QY 251 -----VALVALPROASPLUASPHEVALLEUVALLEUALLEU----- 263
 Db 20532 GCTTTCAGTGTACT-----GACTTCTT-TTAACTGTAACTGTAACTGTAACTTT 20584
 QY 264 -----LeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMet----- 277
 Db 20585 CTTGGGCGACGCTATACACGATCTCTGGAATCTGACAGACAGTGCCTCAAAACCG 20644
 QY 278 -----GLYARGCYSALAALAGLYVALMETHISTLEUPHETYL----- 289
 Db 20645 AGGAGACACCGCCAGGTGTGCA-----CACAGCTTTCTGTCAAGATTTGGGA 20692
 QY 290 -----VALARGALAGLYVA 294
 Db 20693 GGACTCTGGGATCCCTGACACACATCTGTTCATGCGACCTTAGGTAAAGGCTCTGT 20752
 QY 294 1SERARGALAMETLEUARGLEUPHELEUALAMETGLULYSGLYARGHISMETGLUTY 314
 Db 20753 TCAAGGAGGCTTTGCTCTGTGTG-----GGTGGATGGGGAGAGT 20794
 QY 314 RGLU-----CYSPROTYLLEUVA 320
 Db 20795 CTCGAAGCCCTTTRCGGSCCTTGCGATTCCTATMCCCCGGTCTGCGCTGTAGT 20854
 QY 320 1TYVALPROVAL----- 324
 Db 20855 CCAAGTCTCTCTTTTAAACAATGACAGTAATGTACACCGATGACCTTTGGAGACAA 20914
 QY 325 -----VALALAPHEALY-LEUGLUPRO--LYSASPGLYLYSGLYVALPHEALAYLA 341
 Db 20915 TAAAGACCTGATTAATTAATCTTAAACCAACGAGAGAACATCTTTCACGACAC 20974
 QY 341 SPGLYGLUEUMETVALSERGLUALAVALINGLYGLUVALHISPROASNTYRPHETPRM 361
 Db 20975 AACTTCAGTTGATTATGAGCCAAAGGTAAAGAACGCAACGATCTCTTTCG----- 21027
 QY 361 ETVALSERGLYCYSVALGIUPROPROSERTRPLYSPROGLINGMET 377
 Db 21028 -----AAGAAACCTCAGGAGATG 21045
 RESULT 11
 US-09-103-840A-2/c
 ; Sequence 2, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 2436-2007.00
CURRENT APPLICATION NUMBER: US/09/103.840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Alignment Scores:

Pred. No.: 6.08e+04 Length: 4403765
Score: 98.00 Matches: 95
Percent Similarity: 32.75% Conservative: 37
Best Local Similarity: 23.57% Mismatches: 117
Query Match: 4.86% Indels: 156
DB: 4 Gaps: 19

US-09-784-810A-2 (1-384) x US-09-103-840A-2 (1-4403765)

QY 3 Proalaglygylproargglyvalleuproargprocyargvalleuvalleuasn 22
DB 4184549 CCCCTGGCGGACCTCAAGGTGTTCGGAGGGCC 4184514
QY 23 Proargglygylgylgylsalaleuclnleuphearg--Serhizvalglnproleu 41
DB 4184513 -----GGTCGACGACGATCACTCGCTTACGACGTCGCGCGACCGGCT 4184463
QY 42 LeualagluaglualeuSerphenrheumetleuthgluargArgAsnHisAlarg 61
DB 4184462 GACGGTGAAGC-----CATCACGAGACGACCGCATG 4184430
QY 62 GluleuValargSerGluleuGluleuGlyargTTPAspAlaleuValmetSerGlyasp 81
DB 4184429 GAACTG-----ATGCACCGATCGATGGCGT---TTCTTATCACCGAG 4184388
QY 82 GlyleumethisgluValvalasnGlyleumetGluArgproAspTrpGluThalale 101
DB 4184387 TCCGAGAGACATCCCTCATGTCCGCGCGTGCAGAA----- 4184352
QY 102 GlulysProleuCySerleuProalaglySerGlyAsnAlaleuAlaleuSerleuasn 121
DB 4184351 -----CTGTTGACACCCACCGCGGGGGGGGAGCT----- 4184322
QY 122 HisTylalaglyTyrGluGlnValThrasnGluAspLeuThrasnGlyThrleu 141
DB 4184321 -----GGGTTGTGGGGGAACCTATCAG-----CGCATGCTT 4184289
QY 142 LeuCySarGProval-----LeuSerProMetAsnLeuSerleuHisThrAlaser 159
DB 4184288 CAGTCCCGCGAGATGACACCGCTATTTCGCAAGCCCGCCACTTCCTTACCGTCA-- 4184232
QY 160 GlyleuArgSerPheSerValleuSerleuAlatrpGlyPheleAlasrpValaspleu 179
DB 4184231 -----CTCATCACTCGGCTGTGACCGACGCTGACGACCTC 4184190
QY 180 Glu----- 180
DB 4184189 GGCATCACGCGCGCGGTGCTGCGCGCGCGCGCGCGGTGAGCTGCTGGA 4184130
QY 181 -----SerAspIyTyrArgArgLeuGlyGluMetArg 191
DB 4184129 CTGACGTACGGTTACTCTCAATCTGCTCGACCGGACCGCCGCTCTGGGAACACAC 4184070

QY 192 PheThrleuGlyThrPheleuArg-LeuAlaleuAlargThryArgGlyArgleuAl 211
DB 4184069 GTGATCGAGGC-----CTCCGGATGAGACCTTTGGGATATT-----CGAAGATG 4184022
QY 211 aThrleuProvalGlyArgValGlyPheLysThrProalaserProvalvalalGlnG 231
DB 4184021 CACGACGCCCT--GGTCACGCGGGGTGTCGGACTTACGCT-----GA 4183982
QY 231 nGlyProvalaspsalHisleuValProleuGluGlnValProSerHisTrpGlnva 251
DB 4183981 TCCGCCACCGATGACCGACCGAC--CCATGACAGGGCAAGTGGCCGACCGCTGTGACC 4183924
QY 251 lValProAspGluAspPheval----- 258
DB 4183923 GCGGACGACACACCGCCATCAACGACGCGCGCGCGCTGCAACACTGGGGCAT 4183864
QY 259 -----LeuValleuAlaleuHisSerHisleuAl 269
DB 4183863 GCTGGGATCGGTGCGGGGCTGCTCCCTCAACGTTGGGACTGGCGCGTTCGCGTGTAT 4183804
QY 269 aSerGluMetPheAlaAlaPrometGly----- 278
DB 4183803 CGAACACAGATTGACACTGCCCTTGGGGCGCGACACCATCTCAATGTCCCGTGG 4183744
QY 279 -----ArgCysAlaAlaGlyValMetHisleuphetGlyArg----- 291
DB 4183743 GGGACGCGCGCTGCGCGCGCTGCTGCGCGTGGACCGCGGTCAAGCGCGTGAAGA 4183684
QY 292 -----AlaGlyValSerArgAlaMetleuArgPheleuAlametGlyArg 310
DB 4183683 CGCGCGCGGGGTGCAC-----CTCAACGACGTGGGTGGGAGTGGCGCGCGC 4183633
QY 310 gHisMetGluTyr-----GluCySProTyrleuValTyrva 322
DB 4183632 ACTTCGCAATATTCGACGACGACGACGCGCTGCCGCGACGCGCTGTGGCATGT 4183573
QY 322 lProval 324
DB 4183572 TCCGGTG 4183566

RESULT 12

US-09-103-840A-1/C
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 2436-2007.00
CURRENT APPLICATION NUMBER: US/09/103.840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:

Pred. No.: 6.1e+04 Length: 4411529
Score: 98.00 Matches: 95
Percent Similarity: 32.75% Conservative: 37
Best Local Similarity: 23.57% Mismatches: 117
Query Match: 4.86% Indels: 156
DB: 4 Gaps: 19

US-09-784-810A-2 (1-384) x US-09-103-840A-1 (1-4411529)

```

OY      3  ProAlaGlyIleProArgGlyValIleuProArgProCysArgValIleuValIleuAsn 22
Db      4192302  CCCCCTGGCGGAGCCTTCAGAGGTTCGGCGAGGCGCC----- 4192267
OY      23  ProArgGlyIleGlyValIleuValIleuIleuPheArg---SerHisValGlnProIleu 41
Db      4192266  -----GGTCGACGACGATCAGCTCCGCTTACGACGTCCGGCTTCGGCGAGCGGCT 4192216
OY      42  LeuAlaGluAlaGluIleSerPheThrIleuMetIleuThrIleuArgArgSerHisAlaArg 61
Db      4192215  GACGGTGAACG-----CATACGACGAGACACCGCATG 4192183
OY      62  GluIleuValArgSerGluIleuGlyArgTrpAspAlaIleuValIleuMetSerGlyAsp 81
Db      4192182  GAACCT-----ATGTCACCGATCATCGCTG---TTCCTTATCAGCCGAG 4192141
OY      82  GlyLeuMetHisGluValValIleuAsnGlyLeuMetGluArgProAspTrpIleuThrAlaIle 101
Db      4192140  TCGCGAGACATCCGCTGATCGCGCGCTGCAG----- 4192105
OY      102  GlnIleuProIleuCysSerIleuProAlaGlySerGlySerGlnAlaIleuAlaIleuSerIleuAsn 121
Db      4192104  -----CTGTTGAGACCCACCGCGCGCGGAGGAGCT----- 4192075
OY      122  HisTrpAlaGlyIleuGluIleuValIleuAsnGluAspIleuThrAsnGlyThrIleuLeu 141
Db      4192074  -----GGGTTGTGCGGGAACCTATCAG-----GCGATGCTT 4192042
OY      142  LeuCysArgProVal-----LeuSerProMetAsnIleuSerIleuHisThrAlaSer 159
Db      4192041  CAGTGGCGGAGATAGACACCGCTATTTCGACAGCGCCACGCTGCTCAGCGTGCA--- 4191985
OY      160  GlyIleuArgSerPheSerValIleuSerIleuAlaTrpGlyPheIleuAlaIleuValIleu 179
Db      4191984  -----CTCATACACCTCGCGTGGTGGTGCACGACCGACCTACGTCGACCTC 4191943
OY      180  Glu----- 180
Db      4191942  GGGTACACGCGCGGGGTCTGCCCTCGCGCGCGCGGCGGCTGCTGAGAA 4191883
OY      181  -----SerAspIleuTrpArgArgIleuGlyIleuMetArg 191
Db      4191882  CTGACGTCAACGGTACACTCCATCTGCTGCACCGCGACCGCGCTCGGAAACACAC 4191823
OY      192  PheThrIleuGlyIleuPheIleuArg-IleuAlaIleuArgTrpThrArgGlyArgIleuAl 211
Db      4191822  GTGATCGAGGCG-----CTCCGGGATGAGCGCTTTCGATCTAT-----CGAAGATG 4191775
OY      211  aThrIleuProValGlyArgValGlyPheIleuTrpAlaSerProValValIleuGlnG 231
Db      4191774  CACCAAGCCCT--GGTCGACGGGGTGTCCGGAATTACGCT-----GA 4191735
OY      231  nGlyProValAspAlaHisIleuValIleuGluIleuGlnValIleuProSerHisTrpGlnVal 251
Db      4191734  TCGCCGACGCGATGACACACGAC--CCATCGAGGACAGATTGCGACACCGCTGGTCA 4191677
OY      251  lValIleuProAspGluAspPheVal----- 258
Db      4191676  GGGGACGACACACCGGCATAGACGACGCGCGCGCTTCACACAGCTGGGGGCAT 4191617
OY      259  -----LeuValIleuAlaIleuIleuHisSerHisIleuAl 269
Db      4191616  GCTGGGATGCTGGTCCGGGCTCGCTCCCTCAACGTTGCGACTGCGCGCTTCGCGCTGAT 4191557
OY      269  aSerGluMetPheAlaIleuAlaIleuMetGly----- 278
Db      4191556  CGAACAAACAGTTGACCTCCCTTGGGGGCGCGCACACACATGCTCATATGTCGCGTGG 4191497
OY      279  -----ArgGlyAlaAlaGlyValIleuMetHisIleuPheTrpValArg----- 291
Db      4191496  GGGAGCGCGGCGCTGGCGCGGCGACGCTTGGCGCTGACCGGCTCAAGCGCGTGAAGA 4191437
OY      292  -----AlaGlyValSerArgAlaMetIleuIleuArgIleuPheIleuAlaMetGluIleuGlyArg 310

```

```

Db      4191436  CGCGCGCGGGGTGAGC-----CTCAACGACGTGTGTGGCGATGGTGGCGCGCGC 4191386
OY      310  gHisMetGluTrp-----GluCysProTrpIleuValIleuVal 322
Db      4191385  ACTTCGGCAATATCTGAGACACACAGCGCGCTGCCGACACACGCGCTGTTCGATGGT 4191326
OY      322  lProVal 324
Db      4191325  TCCGGTG 4191319

RESULT 13
US-09-773-816-1
; Sequence 1, Application US/09773816
; Patent No. 6340774
; GENERAL INFORMATION:
; APPLICANT: Stanford University
; APPLICANT: Khosla, Chaitan
; TITLE OF INVENTION: NON-STEROIDAL ESTROGEN-RECEPTOR
; FILE REFERENCE: 28600-20210.00
; CURRENT APPLICATION NUMBER: US/09/773,816
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/243,458
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 23673
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(23623)
; OTHER INFORMATION: n = A,T,C or G
US-09-773-816-1

Alignment Scores:
Pred. No.: 24.2 Length: 23673
Score: 96.00 Matches: 94
Percent Similarity: 35.04% Conservative: 50
Best Local Similarity: 22.87% Mismatches: 129
Query Match: 4.76% Indels: 138
Gaps: 21

US-09-784-810a-2 (1-384) x US-09-773-816-1 (1-23673)
OY      6  GlyProArgGlyValIleu-----ProArgProCysArgValIleuValIleuAsn 22
Db      4706  GGGCCCGCGCTCGCGGAGTACGGAACACCCCGCGGCTCGCTGCGCGGG 4765
OY      23  ProArgGlyGlyIleGlyValIleuGlnIleuPheArgSerHisValGlnProIleu 42
Db      4766  CGGGCGGAGGCGGGGTGCGGCTGCTGGAGCACCGCGCGGACCGCGCTTCA 4825
OY      43  AlaGlu-----AlaGluIleSerPheThr 50
Db      4826  GCGGACAGGGGACTCCCGTCCGCGCTTCGCGGCTCGCGGAGACATCTGTGTGACG 4885
OY      51  -----LeuMetIleuTrpIleuArgAsn 58
Db      4886  AGGAGGTGCTGTGGCTCCGCGGCGGAATTGTCTGTGTCGACACCTCGCGACGAAC 4945
OY      59  HisAlaIleuGluValIleuValIleuArgSerGluIleuGlyArgTrpAspAlaIleuMet 78
Db      4946  CCGATG-----GTGGTGGCGCAC----- 4975
OY      79  SerGlyAspIleuMetHisGluValValIleuAsnGlyLeuMetGluArgProAspTrpGlu 98
Db      4976  -----ACCTGAACTGTGTG 4990

```

```

QY 99 ThrAla-----IleGlnLysProLeuCySerLeuProAlaGlySerGlyAsnAlaLeu 116
DB 4991 AGCGGGGGCGGATGCGGGCGAGGGCGCTTTCGGCGTCCGGGCCACACAGCAGCTTG 5050
QY 117 AlaAlaSerLeuAsnHisTyrAlaGlyTyrGlnGlnValThrAsnGlnAspLeuThr 136
DB 5051 GCCACAGCAGAGTCTACGACGCGGGC-----ACGGTGTACCGG 5089
QY 137 AsnCyThrLeuLeuLeuCyArgProValLeuSerPrometAsnLeu-----LeuSerLeu 155
DB 5090 GCGTCGACGTGTGTCTGACGCGGAGAGCGGGCGCGGGCGGTGAACCTGTCTGATC 5149
QY 156 HisThrAlaSerGlyLeuArgSerPheSerValLeuSerLeuAlaTropolyPheIleAla 175
DB 5150 AGCCTGGCGCTCGGACGAAATCCG-----5173
QY 176 AspValAspLeuGlnSerAspLysTyrArgArgLeuGlyLys-----MetArgPhe 192
DB 5174 -----CGCGTCGGGTCTCTCGCGCTGTATGCGGCAC 5203
QY 193 ThrLeuGlyThrPheLeu---ArgLeuAlaAlaLeuAlaGlyThrTyrArgGlyArgLeuAla 211
DB 5204 TCCATGAGACACGCGCTGTGAGTGTGATGCTTCTGCGCGACCTGACGCGAGCGCGGGG 5263
QY 212 ThrLeuProValGlyArgValGlyPheLeuThrProAlaSerProValValAlaGlnGln 231
DB 5264 GCGAC-----AGGATCTGCTCGCGCACCCAGGTGATCCGAGTGCATCTCTGTG 5314
QY 232 GlyProValAspAla-----HisLeuValProLeuGlnGlnValProSerHis 248
DB 5315 ACCGGGTCTGACCTGATACGGGACGAGTTCACCTTCATGAGAGAACGATTCCTCGCGGC 5374
QY 249 TrpGlnValValProAspGlnAspPheValLeuValLeuLeuLeuHisSerHisLeu 268
DB 5375 -----GCGAGAGAACTCTGAACGCTG 5395
QY 269 -----AlaSerGlyMetPheAla 274
DB 5396 CCGGCGCGAGCTAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5455
QY 275 Ala---PrometGlyArgCysAlaAlaGlyValMetHisLeu---PheTyrValArgAla 292
DB 5456 GCGAGCGCGCTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5515
QY 293 GlyValSerArgAlaMetLeuLeuArgLeuPhe---LeuAlaMetGlyCysGlyArgHisM 312
DB 5516 TGCACGAGCACTCCGCTTCCGCCACAGTGTGACGAGCTTCCGTACCGGTGCGCGAGCACC 5575
QY 312 etGlnTyrGlnCyArgProTyrLeu-----ValTyrValProValAla 326
DB 5576 TGGATCTGAGCTGCGCGGGCGCTCGGGCAGGTAGCGCTGACGTACACCGCGCGCGCGCG 5634
QY 326 LaphArgLeuGlnProLysAspGlyLysGlyValPheAlaValAspGlyGlyLeuMetV 346
DB 5635 -----GAACACCGCCAAACGCGGTGCTGCTGCGCGGAA-----CGTG 5674
QY 346 alSerGlnAlaValGlnGlyGlnValHis 355
DB 5675 TCGGGGAAGTCTCTCGCTGAGTGTGACAC 5703

```

RESULT 14

US-08-923-137-2/c

Sequence 2, Application US/08923137

Patent No. 6083716

GENERAL INFORMATION:

APPLICANT: Wilson, James M.

APPLICANT: Farina, Steven F.

APPLICANT: Fisher, Krishna J.

TITLE OF INVENTION: Chimpanzee Adenovirus Vectors

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr., P.O. Box 457

```

CITY: Spring House
STATE: Pennsylvania
COUNTRY: United States of America
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,137
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/024,700
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVN.02ICPIUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 36519 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-923-137-2

```

```

Alignment Scores:
Pred. No.: 63.4 Length: 36519
Score: 95.00 Matches: 81
Percent Similarity: 34.16% Conservative: 43
Best Local Similarity: 22.31% Mismatches: 133
Query Match: 4.71% Indels: 106
DB: Gaps: 17

```

```

US-09-784-810A-2 (1-384) x US-08-923-137-2 (1-36519)
QY 56 ArgArgAsnHisAla-----ArgGlnLeuValArgSerGlnGlnLeuGlnArgTyrPae 73
DB 12717 CCGCGTACAGCAGCTGTGACAGCAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCG 12658
QY 74 AlaLeuValAlaMetSerGlyAspGlyLeuMetHis-----GlnValValAsnGly 90
DB 12657 GCCCTGCAAGTCTTGAAGACCTGTGCTGACAGCTTCAAGGCTGTCAAGCCCTGCGACTGGT 12598
QY 91 LeuMetGln-----ArgProAspTropolyThrAlaIle-----101
DB 12597 CTGGAAGAAGTACTCCGCGCGCGAGCTGTGACACTGTGACACTGTGCGGTGCGGCACCTCGGTACCAT 12538
QY 102 ---GlnLysProLeuCySerLeu-----ProAlaGlySer-----112
DB 12537 CAGGGCGAGGGCGCTGATGAAGTGTGTGATGCTCTCTGCGCGCGCGCGCGCGCGCG 12478
QY 113 -----GlyAsnAlaLeuAlaAspLeuAsn-----121
DB 12477 CTGGGTGTGAGGAAGCGCTTCAAGCGCCACATGATCCACAGTTGAGTCCGCGGTGA 12418
QY 122 -----HisTyrAlaGlyTyrGlnGlnValThrAsnGlnAspLeuLeuThrAs 137
DB 12417 CCGCTCGCGCTGGCGCCAGCGCTTCCGCGCACGTCGCTGATCCATGCGGTCCAGGTGATC-- 12360
QY 137 nCyThrLeuLeuLeu-----CysArgPr 145
DB 12359 -TGCAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 12301
QY 145 oValLeuSerPro---MetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPh 164
DB 12300 GGCCTGTCTCGCGCGCGGATGCGCTGTGTCTTCTACACAGCGGCTTTCAGATGCGCAGGACCT 12241

```

QY 164 eserValleuSerleuAlaTrpGlyPheleAlaAspValaAspLeuGluSerAspLysTy 184
DB 12240 CTGCTGGGTGGGGTGGT-----GACCGCGAGAGGACACAGCGC 12205
QY 184 rArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeuAlaAlaLeuAr 204
DB 12204 CTCAGAGATGGCCGAGAGCGGTGGCTGGCGCTGGCTGTAAGAGCTTCGGGGTTCGG 12145
QY 204 gThrTyArgGlyArgLeuAlaThrLeuProValaGlyArgValaGlyPheLysThrProAl 224
DB 12144 GGTGTCAGCGGCATGATGCTTGCATG-----12117
QY 224 aserProvalValaGlnGlnGlyProValaAspAlaHisLeuValProLeuGluGlu 244
DB 12116 -----GCTGGGTGCATGCTCGAGAGAGTT 12091
QY 244 nValProSerHisTrp-GlnValValPro-----AspGluAspPheValLeuV 260
DB 12090 AATCCCGACGCGCTGCTGTCAGCGCGCCGCGCATCGCGGATCAGAGAGTGGCTTGG 12031
QY 260 alLeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgC 280
DB 12030 TTGTTCATCTACCAAAATACGTTCGCGCATCATCTTCAGAGTA--CTCGCCCTCC 11974
QY 280 ySaLaAlaGlyValaMetHisLeuPheTyValaArgAlaGlyValaSerArgAlaMetLeuL 300
DB 11973 TCGTCCACCTCATCGTCCACCTCTCTACGTA--GGTCTCGCGCGCGCGCTCCAGAGCC 11917
QY 300 euArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyArgLysProTyLeuV 320
DB 11916 CGCGGCTGGGTGGCCAGTGCAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11876
QY 320 aTyValProValaAlaPheArgLeuGluProLysAspGlyLysGlyValaPheAlaV 340
DB 11875 -----GTAAGTCTCCCTCCCTGCTGCTCGCGCGCGCGCGCGCGCGCGCGCGCA 11830
QY 340 alAspGlyGluLeuMetValaSerGlu-----AlaValGlnGlyGlnVal 355
DB 11829 CTAGCATCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11770
QY 355 lSPro 356
DB 11769 ATCCCT 11765
RESULT 15
US-08-845-566-2/C
Sequence 2, Application US/08845566
Patent No. 591214
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl
TITLE OF INVENTION: EDG-1 LIKE RECEPTOR
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,566
FILING DATE: Filed Herewith
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0271 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1649 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TYNMOR01
CLONE: 144690
US-08-845-566-2
Alignment Scores:
Pred. No.: 0.57 Length: 1649
Score: 94.00 Matches: 83
Percent Similarity: 32.08% Conservative: 36
Best Local Similarity: 22.37% Mismatches: 137
Query Match: 4.67% Indels: 20
Gaps: 20
US-09-784-810A-2 (1-384) x US-08-845-566-2 (1-1649)
QY 1 MetAspProAlaGlyGlyPro-----7
DB 1612 TTGACACCTTTAAAGNCCCGCCCGCTTTNTTACACCTGANGNTTTCACGNTACC 1553
QY 8 -----ArgGlyValLeuProArg-----ProCysArgVal 17
DB 1552 ATTATATCAAACTCAATGTTCCAGAGGACATGATCCCAAGTCCACCCCTTGACGCTG 1493
QY 18 LeuValLeuLeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHis 37
DB 1492 GGACCCCTACTGAGGCGCCAGAGAGGCGCGGA-----CTTCCCAAGATATAC 1445
QY 38 ValGlnProLeuLeu-AlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgAr 57
DB 1444 ACAGAACCGGGGTGTGGGGAA-----CCCACTCG 1415
QY 57 GAsnHisAlaArgGluLeuValaArgSerGluLeuGlyArgTyTPAspAlaLeuVala 77
DB 1414 CCCCCACACACG-----GTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1361
QY 77 lMetSerGlyAspGlyLeuMetHisGluValaValaAsnGlyLeuMetGluArgProAspTr 97
DB 1360 TGTCCAGAGTGAC-----CATGG-----AACCCATTGG 1331
QY 97 pGluThrAlaIleGlnLysPro-----LeuCysSerLe 108
DB 1330 GCGT-----CAGAACCCCGAGAGGCCACCGGAGAACATGAGGAGCTGTCCCT 1280
QY 108 uProAlaGlySerGlyAsn-----114
DB 1279 TCCCTGCTCCCATACAGCGAGGCTGTCACACAGCTTCCTGACCGAGAGGAC 1220
QY 115 -----AlaLeuAlaSerLeuAsnHisTyAlaGlyTyArgGluValaThrAsn-- 131
DB 1219 CTGCTGCGACAGCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1160
QY 132 -----GluAspLeuLeuThrAsnGlyThrLeuLeuLeuGlyArgProValaLeuSerPr 149
DB 1159 CCGACAGCTGAGATGCTGAGACAGGGGCTCCCGCATCCGAAACCTGACGACGAGGAGACC 1100
QY 149 oMetAsnLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSerLe 169
DB 1099 GCGAAGCTG-----TCCCTTGCTCAGAGAGCTGCGTGGTGGAG 1055
QY 169 uAlaTrpGlyPheAlaAsp-----ValaAspLeuGlu 180

```
Db 1054 CTCC---GAGTGAAGCTTCAGAGGCCGCGGCCAGAGTCCCGCGGCTTCGCATGGCC 998
QY 180 uSerAspIysTyrTrgArgLeuGIyGIuMetArgPheThrLeuGIYThrPheLeu----- 198
Db 997 AGCGGAGACACCCGACAGAGGAGGCTGAGCAGCGGCTCTGCACACCTCCCTGTCGCG 938
QY 199 -----ArgLeuAlaAlaLeuArgThrTyrTrgGIYArgLeuAlaTh 212
Db 937 GAAGAGTAgATGATGGGGTTGACCGCCGAGTTGAGAGCGGCCAGGGCCAGGATCCAGTC 878
QY 212 rLeuPro-----ValGIYArgValGIYpHeLysThrProAlaSerProVa 227
Db 877 CATGCCCGCAGTACTCTGCTGGGCCAGAGGCTTGAGCCAAAGAGTCGCGCCAGC----- 823
QY 227 lValValGIuGIuInGIYProValAspAlaHisLeuValProLeuGIuGIuInValProSe 247
Db 822 -----AGCAGCCCC-----GAAAGTGGGCCCCAGC 797
QY 247 rHisTrpGIuValValProAspGIuAspPheValLeuValLeuAlaLeuHisSerHI 267
Db 796 ACACGAGAGAGGCCAGCAGAT-----CATCAGCA 767
QY 267 sLeuAlaSerGIuMetPheAlaAlaProMetGIYArgCysAlaAlaGIYValMetHisLe 287
Db 766 CCGTCTTCAGCAGCGCGGCGGCTTGCAGCGGCGCTGCGGCGCTGCGGCGCTTCTG----- 712
QY 287 uPheTyrValArgAlaGIYValSerArgAla 297
Db 711 -----CCGCTGGCTTCACACAGCG 691
```

Search completed: July 12, 2003, 20:38:12
Job time : 3167 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:08:03 ; Search time 293 Seconds

(without alignments)
2951.424 Million cell updates/sec

Title: US-09-784-810a-2

Perfect score: 2015

Sequence: 1 MDPAGGPRGVLPKRCRLVL.....CVEPPPSMKPQOMPPEEPL 384

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Command line parameters:
-MODEL=frame+pn.model -DEV=xlh
-Q=/cgn2.1/USPTO_SPOOL/US09784810/runat.11072003_103441_26547/app_query.fasta.1.583
-DB=N.Geneseq.101002 -QEXT=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09784810.ecgn1.1.263.etrnat.11072003.103441.26547 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N.Geneseq.101002.*
1: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2015	100.0	1600	22	AAD14424	Human sphingosine
2	1962	97.4	1573	21	AAV75676	Human encoding a hum
3	1962	97.4	1719	22	AAD04477	Human sphingosine
4	1962	97.4	1772	22	AAH16415	Human cDNA sequenc
5	1962	97.4	1821	22	AAH15652	Human cDNA sequenc
6	1959	97.2	1173	24	ABL59533	Human sphingosine
7	1959	97.2	1205	22	AAAC84161	Human sphingosine
8	1957	97.1	1438	22	AAI59336	Human polynucleoti
9	1946	96.6	1447	21	AAAS0508	Human polynucleoti
10	1944	96.5	1562	24	ABL90618	Human polynucleoti
11	1927	95.6	1533	20	AAV84490	Human secreted pro
12	1927	95.6	1533	22	ABA83273	Human secreted pro
13	1594.5	79.1	1759	22	AAD14425	Mouse consensus sp
14	1585.5	78.7	1559	21	AAZ47168	Mouse sphingosine
15	1582.5	78.5	1815	21	AAZ47167	Mouse sphingosine
16	1557	77.3	1200	22	AAI61122	Human polynucleoti
17	875	43.4	2698	22	AAI51816	Human polynucleoti
18	856.5	42.5	2380	22	AAI51817	Human cDNA for sph
19	856.5	42.5	2380	24	ABL59533	Human cDNA for sph
20	839.5	41.7	2422	23	AAH85331	Human sphingosine
21	673.5	33.4	11096	22	AAI18278	DNA encoding novel
22	550	27.3	2020	23	ABL03297	Human nervous syst
23	550	27.3	4020	23	ABL03296	Drosophila melanog
24	493.5	24.5	1578	23	AAH85330	DNA encoding novel
25	492.5	24.4	2830	23	ABL02327	Drosophila melanog
26	467	23.2	2136	22	AAK94172	Human full-length
27	448	22.2	587	22	AAH10443	Human cDNA clone (
28	447.5	22.2	675	22	AAH07214	Human cDNA clone (
29	436	21.6	801	22	AAH99234	Human protein enco
30	427.5	21.2	5886	23	ABL02326	Drosophila melanog
31	403	20.0	240	22	AAD04478	Conserved region D
32	365	18.1	296	22	AAAS24477	Human ovarian PCR
33	365	18.1	296	22	AAH83083	Human ovarian tumo
34	339	16.8	604	22	AAH90477	Human cDNA clone (
35	319	15.8	539	22	AAK92493	Human cDNA 3'-end
36	301.5	15.0	1521	22	AAD08644	Human kinase (PKIN
37	300	14.9	1394	24	ABO54324	Human ovarian anti
38	290	14.4	297	24	ABN22070	Human ORFX polynuc
39	268.5	13.3	4413	24	ABL40828	Human sphingosine
40	266.5	13.2	4231	21	AAAS0510	Human sphingosine
41	260.5	12.9	4463	24	ABA96945	Human ceramide kin
42	255	12.7	1840	22	AAD14426	Human sphingosine
43	233	11.6	2173	23	ABL25705	Drosophila melanog
44	233	11.6	10337	23	ABL25704	Drosophila melanog
45	228	11.3	394	24	ABL66446	Lung cancer relate

ALIGNMENTS

RESULT 1	
AAD14424	
ID	AAD14424 Standard; cDNA; 1600 BP.
XX	
AC	AAD14424;
XX	
DT	01-NOV-2001 (first entry)
XX	
DE	Human sphingosine kinase (Spk) cDNA #1.
XX	
KW	Human; sphingosine kinase; Spk; restenosis; ischaemia; gene therapy;
KW	antitense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
KW	cytostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
KW	leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.
XX	
OS	Homo sapiens.
XX	
FX	Key
XX	Location/Qualifiers

```
FT      CDS      176..1330
FT      /tag- a
FT      /product- "Human sphingosine kinase (Sphk) protein #1"
XX      MO200160990-A2.
XX      23-AUG-2001.
XX      14-FEB-2001; 2001WO-US04789.
XX      14-FEB-2000; 2000US-0182360.
XX      22-MAR-2000; 2000US-0191261.
XX      (CURA-) CURAGEN CORP.
XX      (GETH ) GENENTECH INC.
XX      Rastelli L;
XX      WPI; 2001-514770/56.
XX      P-PSDB; AAE07882.
XX      An isolated sphingosine kinase polypeptide useful for treating a
XX      Sphk-associated disorder especially cancer, restenosis or ischemia in a
XX      human -
XX      Claim 8; Page 89-90; 107pp; English.
XX      The present invention relates to sphingosine kinase (Sphk) polypeptides
XX      and nucleic acids encoding them. Sphk is useful for treating a Sphk-
XX      associated disorder especially cancers such as leukaemia, lymphoma,
XX      ovarian, breast, lung, colon, testicular, stomach and skin,
XX      atherosclerosis, restenosis or ischaemia and cell proliferative disease
XX      or disorder associated with vascular diseases. Sphk gene is used in gene
XX      therapy and antisense-therapy. Sphingolipids serving as signalling
XX      molecules, have recently emerged as regulators of cell growth,
XX      differentiation, diverse cell phenotypes and cell death. Activation of
XX      Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
XX      endothelial cells. The present sequence is human sphingosine kinase
XX      (Sphk) cDNA.
XX      Sequence 1600 BP; 265 A; 492 C; 531 G; 311 T; 1 other:
XX      Alignment Scores:
XX      Score: 6,24e-194 Length: 1600
XX      Percent Similarity: 2015.00 Matches: 384
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 22 Gaps: 0
XX      US-09-784-810A-2 (1-384) x AAD14424 (1-1600)
XX      1 MetAspProAlaGlyValGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
XX      176 ATGGATCCAGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 235
XX      21 LeuAsnProArgGlyValGlyValGlyValGlyValGlyValGlyValGlyValGlyValGly 40
XX      236 CTGAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 295
XX      41 LeuLeuAlaGluAlaGluLeuLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
XX      296 CTTTGGCTGAGGCTGAATCTCTCTACGCTGATCTCTACGCTGAGCGGCGGGAACCAACGCGG 355
XX      61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValValMetSerGly 80
XX      356 CGGGAGCTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 415
XX      81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
XX      416 GACGGGCTGTGATGACGAGGCTGTGACGAGGCGGCTGATGAGCGGCGGCGGCGGCGGCGGCGG 475
XX      101 IleGluInsProLeuInsSerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu 120
```

```
DB      476 ATCCAGAGGCCCTCTGTAGGCTCCAGCAGGCTCTGGCAACGGCGTGGCAGCTTCCTTG 535
XX      121 AsnHisTyrAlaGlyTyrGluGluValThrAsnGluAspLeuThrAsnGlyThrLeu 140
XX      536 AACCATATGCTGGCTATGACAGCAGTCAACCATATGACATGACCTCTGACCACTGACGCTA 595
XX      141 LeuLeuGlyArgProValLeuSerPheMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
XX      596 TTGCTGTCCGCGCGGCTCTCTACCCATGAACTGCTGTCTGCAACAGCGGCTTCGGGG 655
XX      161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu 180
XX      656 CTGGCGCTGCTCTCTGCTCAGCTCAGCTGCGGCTGATGCTGATGCTGATGCTGATGCTGATG 715
XX      181 SerAspLysTyrArgArgLeuGlyValMetArgPheThrLeuGlyThrPheLeuArgLeu 200
XX      716 AGTGATTAAGTATGCGGCTGTGGGGAGATGCGCTTCACTCTGCGGCACTTCTCTGCTG 775
XX      201 AlaAlaLeuArgThrTyrArgGlyValGluAlaThrLeuProValGlyArgValGlyPhe 220
XX      776 GCAGCCCTGCGCAGCTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 835
XX      221 LysThrProAlaSerProValValValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
XX      836 AAGACACCTGCTTCCCGCTGTGTGTCAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 895
XX      241 LeuGluGluGluValProSerHisTyrGlnValValProAspGluAspPheValLeuVal 260
XX      896 CTGGAGAGCAGCAGTGCCTCTCTACGTCAGCGTGGTGGCGGAGAGAGCTTGTGCTAGTGC 955
XX      261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280
XX      956 CTGGCAGCTCTGTCACACCGCAGCTGGCAGTGAAGTGTGCTGCAACCATGGCGGCTGT 1015
XX      281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu 300
XX      1016 GCAGTGGCGCTATGATGATGCTTCTACGTCGCGGCGGAGTGTCTGTCGTCATGCTGCTG 1075
XX      301 ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProThrLeuVal 320
XX      1076 CGCCTCTCTCTGCGCATGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1135
XX      321 TyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal 340
XX      1136 TATGTGCGCGCTGTGCTGCGCTCTGCGCTGAGACCCCAAGAGTGTGTTGCGAGTG 1195
XX      341 AspGlyLeuLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrp 360
XX      1196 GATGGGAATGTGATGTTAGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1255
XX      361 MetValSerGlyCysValGluProProPheSerTrpLysProGlnGlnMetPropProPro 380
XX      1256 ATGTGTACAGCGTGTGCGGAGCCCGCGCGAGCTGGAAGCCCAAGCATGTCACCGCGCA 1315
XX      381 GluGluProLeu 384
XX      1316 GAAAGGCCCTTA 1327
XX      DB
XX      RESULT 2
XX      ID AAA75676 standard; DNA: 1573 BP.
XX      AC AAA75676;
XX      DE 22-JAN-2001 (first entry)
XX      DNA encoding a human regulator of intracellular phosphorylation.
XX      Human: intracellular phosphorylation regulator; HRIP; stroke; myeloma;
XX      neurological disorder; Parkinson's disease; demyelinating disease;
XX      meningitis; developmental disorder; neuromuscular disorder; cancer;
XX      myasthenia gravis; cell proliferative disorder; actinic keratosis;
```

KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;
 KW autoimmune disorder; inflammatory disorder; Addison's disease;
 KW acquired immunodeficiency disease; allergy; diabetes mellitus;
 KW rheumatoid arthritis; microbial infection; trauma; ss.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 130..1284
 FT /*tag= a
 FT /product= "regulator of intracellular phosphorylation"

XX WO20005332-A2.

XX 21-SEP-2000.

XX 17-MAR-2000; 2000WO-US07277.

XX 18-MAR-1999; 99US-0125593.

XX 20-MAY-1999; 99US-0135049.

XX 09-JUL-1999; 99US-0143188.

XX (INCY-) INCYTE PHARM INC.

XX Baugman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
 PI Lu DAM, Au-Young J;

XX WPI: 2000-602121/57.

XX P-PSDB: MAB18659.

XX Novel human intracellular phosphorylation regulator polypeptides and
 PT polynucleotides for diagnosis, prevention and treatment of
 PT neurological, cell proliferative and autoimmune/inflammatory disorders

XX Claim 4; Page 89; 96pp; English.

XX The present sequence encodes a human regulator of intracellular
 CC phosphorylation (HRIP). HRIP is useful for screening agonists and
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
 CC are useful for treating a disease or condition associated with
 CC decreased or increased expression of functional HRIP. Diseases treated
 CC or diagnosed include neurological disorders such as stroke, Parkinson's
 CC disease, demyelinating diseases, bacterial and viral meningitis and
 CC other developmental disorders of the central nervous system.
 CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
 CC inflammatory disorder such as Addison's disease, acquired
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
 CC rheumatoid arthritis, microbial infection and trauma.

XX Sequence 1573 BP; 274 A; 480 C; 514 G; 305 T; 0 other;

XX Alignment Scores:

Pred. No.: 1,41e-188 Length: 1573
 Score: 1962.00 Matches: 375
 Percent Similarity: 98.18% Conservative: 2
 Best Local Similarity: 97.66% Mismatches: 7
 Query Match: 97.37% Indels: 0
 DB: 21 Gaps: 0

US-09-784-810A-2 (1-384) x AAA75676 (1-1573)

OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
 DB 130 ATGGATCCAGCGGGGGCCCGGGGGCGTCCCGGGCCCTGCGCGCTGAGTCTG 189
 OY 21 LeuAsnProArgGlyGlyGlyGlyAlaLeuGlnLeuPheArgSerHisValGlnPro 40
 DB 190 CTGAACCCGCGGGGGCAAGGGCAAGCCCTTGACAGCTTCCGAGTACGTGAGCCC 249

OY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
 DB 250 CTTTGGCTGAGAGCTGAATCTCTTACAGCTGATGTCACAGCGGAGCAACACCGC 309
 OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTTPAspAlaLeuValMetSerGly 80
 DB 310 CGGGAGCTGTGCGGTCCGAGAGAGCTGGGCCCTGGAGCGCTGTGGTCAATCTGGA 369
 OY 81 AspGlyLeuMetHisGluValAlaValAsnGlyLeuMetGluArgProAspTProGluHisAla 100
 DB 370 GACGGCTGATGACAGAGTGTGTGACAGGGCTCATGAGAGCGCTGACTGGAGAGCCGCC 429
 OY 101 IleGlnIysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
 DB 430 ATCCAGAAAGCCCGTGTAGCCTCCAGAGGCTGTGGAAAGCGCTGGACCTTCCCTTG 489
 OY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeu 140
 DB 490 AACCATTTATGCTGCTATGACAGAGCTACCAATGAAGACCTTCGACCAACTGACGCTA 549
 OY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
 DB 550 TTGCTGTGCGCGCGCTGCTGTACACCATGAACCTGCTGTCTGTGACACGCGCTTCCGGG 609
 OY 161 LeuArgSerPheSerValLeuSerLeuAlaTrrGlyPheIleAlaAspValAspLeuGlu 180
 DB 610 CTGGCCCTTCTCTGTGCTGACCTGGCGCTGGGCTTCATCTGTGATGTGACCTAGAG 669
 OY 181 SerAspIysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyTyrPheLeuArgLeu 200
 DB 670 AGTGGAAGTATGCGCGCTGTGGGAGATGCGCTTCACCTGCGGACCTTCTGCGGTG 729
 OY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
 DB 730 GCAGCCCTGCGCAGCTACCGCGCGAGCTGAGCTTCACTTCTGTGAGAAAGATGGCTTCC 789
 OY 221 LysThrProAlaSerProValAlaValGlnGlnGlyProValAspAlaHisLeuValPro 240
 DB 790 AAGACACCTGCTCCCGCTTGTGTGTCACGAGGCGCGGTGATGACACCTTGTGCA 849
 OY 241 LeuGluGluGlnValProSerHisTrrGlnValValProAspGluAspPheValLeuVal 260
 DB 850 CTGGAGGAGCCAGTCCCTCTCATGAGAGTGGTCCGAGAGAGACTTGTGCTAGTC 909
 OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280
 DB 910 CTGGACACTGCTGACCTGCGACCTGGGACAGTGTGCTGACCCCATGGCGGCTGT 969
 OY 281 AlaAlaGlyValMetHisLeuPheTyrValAlaGlnGlyValSerArgAlaMetLeuLeu 300
 DB 970 GCAGCTGGCGTATGATCTGTTCACGTGGCGGGGAGTGTCTGTCCTCATGCTGCTG 1029
 OY 301 ArgLeuPheLeuAlaMetGluIlysglyArgHisMetGluTyrGluCysProTyrLeuVal 320
 DB 1030 CCCCTTCTCTGCGCATGAGAGGAGGAGGATATGAGTATGAGTATGAGTATGAGTATG 1089
 OY 321 TyrValProValValAlaPheArgLeuGluIlyProIlyAspGlyValPheValAlaVal 340
 DB 1090 TATGTGCCGTGTGTGCTTCCGCTTGAGGCCCAAGAGGAGAGGAGTGTGTTGCACTG 1149
 OY 341 AspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrr 360
 DB 1150 GATGGGAAATTGATGGTTAGCGAGGCGCGTGCAGGGGCGAGGTGCACCAAACTACTTCTG 1209
 OY 361 MetValSerGlyCysValGluIlyProIlyProIlyProIlyPlyPProGlnGlnMetProPro 380
 DB 1210 ATGGTCAGCGGTGTGTGTGAGGCCCGCCGACAGTGAAGCCACAGATGTCACAGCCCA 1269
 OY 381 GluGluProLeu 384
 DB 1270 GAAAGACCTTA 1281

RESULT 3

AD0477	
ID	AD04477 standard; cDNA: 1719 BP.
AC	AD04477;
XX	
DT	04-JUL-2001 (first entry)
XX	
DE	Human sphingosine kinase type I (hsk1) cDNA.
XX	
KM	Human: sphingosine kinase type I; sk1; chromosome 17q25.2;
KM	sphingosine-1-phosphate; S1P; drug screening; therapy; haemostasis;
KM	thrombosis; allergic reaction; proliferative disease; cancer;
KM	haematopoietic disorder; leukaemia; cardiovascular disease; stroke;
KM	atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
KM	autoimmune disease; inflammatory disease; multiple sclerosis;
KM	T helper-1 related disease; chronic obstructive pulmonary disease;
KM	asthma; myocardial infarction; neurodegenerative disorder;
KM	wound healing; embryogenesis; anticoagulant; cerebroprotective;
KM	neuroprotective; antipsoriatic; antiarthritic; cyostatic; cardiant;
KM	vulnerary; ss.
OS	Homo sapiens.
XX	
PH	Key
FT	Location/Qualifiers
CDS	270..1424
FT	/*tag= a
FT	/product= "Human sphingosine kinase type I (hsk1)"
FT	/note= "CDS is specifically claimed in claim 2 and shown as SEQ ID NO 2"
FT	polyA_signal
FT	1675..1681
FT	/*tag= b
FT	misc_feature
FT	264..273
FT	/*tag= c
FT	/note= "translational initiator AUG is in a partial kozak consensus"
PN	WO20011029-A2.
XX	
XX	03-MAY-2001.
PD	
PF	27-OCT-2000; 2000MO-EP09498.
XX	
PR	28-OCT-1999; 99US-0162307.
PR	07-FEB-2000; 2000US-0180525.
XX	
PA	(WARN) WARNER LAMBERT CO.
PX	
PI	Allen J, Gosink M, Melendez AJ, Takacs L;
XX	
DR	WPI: 2001-300510/31.
XX	
PT	P-PsDB; AAE00924.
PT	
XX	
PS	New human sphingosine kinase type I gene for screening drug candidates particularly inhibitors used for preventing or treating e.g. atherosclerosis, thrombosis, asthma and diabetes -
XX	
XX	Claim 2; Fig 1; 91pp; English.
XX	
CC	The present sequence is human sphingosine kinase type I (hsk1) cDNA.
CC	The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the
CC	substrate sphingosine to sphingosine-1-phosphate (S1P). The sk1 gene
CC	and encoded polypeptide are applicable in screening drug candidates
CC	particularly inhibitors for preventing or treating disorders such as
CC	haemostasis, thrombosis, allergic reactions, proliferative diseases
CC	including cancer, haematopoietic disorders such as leukaemia,
CC	cardiovascular diseases such as stroke, atherosclerosis and coronary
CC	artery disease, dyslipidaemia, diabetes including type I and type II
CC	diabetes, autoimmune and inflammatory diseases such as multiple
CC	sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
CC	disease, asthma, myocardial infarction, neurodegenerative disorders,
CC	natural wound healing processes and embryogenesis.
XX	
Sequence	1719 BP; 529 A; 556 G; 329 T; 0 other;

[illegible]

DB 1170 GCCTCTCTGCGCCATGAGAGGAGGAGGAGCATATGAGTATGATGCCCCCTACTGTGTA 1229
OY 321 TyrValProValValAlaPheArgLeuGluProIlysAspGlyValAlaPheAlaVal 340
DB 1220 TATGTGCCCCGTGTCGCTTCCGTGGAGCCCAAGATGGTGTGTTCACGTG 1289
OY 341 AspGlyLeuMetValSerGluAlaValGlnGlyValHisProAsnTyrPheTyr 360
DB 1290 GATGGGAAATTGATGATTAGGAGGAGCCGTGACAGGCCCGACGCCCAACTACTCTCGG 1349
OY 361 MetValSerGlyCysValGluProProSerTyrIlysProGlnMetProPro 380
DB 1350 ATGCTCAGCGGCTTCCTGGAGCCCGCCAGCTGAGAGCCCGACAGATGCCACCGCA 1409
OY 381 GluGluProLeu 384
DB 1410 GAAGAGCCCTTA 1421
RESULT 4
AAH16415
ID AAH16415 standard; cDNA; 1772 BP.
AC AAH16415;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:15393.
XX
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
PD 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Claim 8; SEQ ID 15393; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX oligonucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 1772 BP: 285 A; 566 C; 585 G; 336 T; 0 other.
SQ
Alignment Scores:
Pred. No.: 1.67e-188 Length: 1772
Score: 1962.00 Matches: 375
Percent Similarity: 98.18% Conservative: 2
Best Local Similarity: 97.66% Mismatches: 7
Query Match: 97.37% Indels: 0
DB: 22 Gaps: 0
US-09-784-810A-2 (1-384) x AAH16415 (1-1772)
OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 350 ATGATCCAGGCGGGGCGCCCGGGGCGTGTCCCGCGCGCTGCGCGTGTGCTG 409
OY 21 LeuAsnProArgGlyGlyGlyValAlaLeuGlnLeuPheArgSerHisValGlnPro 40
DB 410 CTGAACCCGCGCGCGGCGGCAAGGCGCCTTGACAGCTCTCCGAGTGCAGCGCC 469
OY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
DB 470 CTTTGGCTGAGGCTGAATCTCTTCACGCTGATGCTCAGTGAAGCGCGGACACACGCG 529
OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTArgPArgAlaValMetSerGly 80
DB 530 CGGAGCTGGTGGCTCGGAGAGCTGGGCGCGTGGGAGCTGTGTGTATGTGTGA 589
OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluValArgProAspTyrGluAla 100
DB 590 GACGGCTGATGACAGAGAGTGTGAACGGCTCATGAGCGCGCTGAGTGGAGAGCGCC 649
OY 101 IleGlnIysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
DB 650 ATCCAGAGCCCTGTGTGATGCTCCACAGAGCTGTGCAAGCGCGCTGAGCTCCCTTG 709
OY 121 AsnHisTyrAlaGlyTyrGluGlnValAlaHisAsnGluAspLeuLeuThrAsnCysThrLeu 140
DB 710 AACCATTTATGCTGTGATAGACAGGTACCAATGAAGACCTCTGACCAACCTGACGCTA 769
OY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
DB 770 TTGCTGTCCCGCGCGCTGCTGTCAACCATGAACCTGTGTGTGTGTGTGTGTGTGTGTG 829
OY 161 LeuArgSerPheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeuGlu 180
DB 830 CTGCGCCCTCTCTG 889
OY 181 SerAspIysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
DB 890 AGTGAAGATATGCGCGTGTGGGAGATGCGCTTCACTGTGGGACCTTCTGTGTGTGTG 949
OY 201 AlaAlaLeuAspTyrThrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
DB 950 GCAGCGCTGCGGACCTACCGCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1009
OY 221 LysThrProAlaSerProValValAlaGlnGlnGlyProValAspAlaHisLeuValPro 240
DB 1010 AAGACACCTGCTCCCGCGTG 1069
OY 241 LeuGluGluGlnValProSerHisThrPheValValProArgGluAspPheValLeuVal 260
DB 1070 CTGAGGAGCCAGTGTCTCTTCACTGTGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1129
OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280

Db 679 AAGACACCTGCTCCCGCTTGTGGTCCAGACGAGCCGCGTATGACACCTTGTGCCA 738
 Qy 241 LeuGlulGluValProSerHisTrpGlnValProAspGluAspPheValLeuVal 260
 Db 739 CTGGAGAGACCAAGTCCCTCTCAGTGCAGACAGTGGTGCCGACGAGACTTGTGTACTC 798
 Qy 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280
 Db 799 CTGGCACTGCTGCACCTGCAGACCTGGGCACTGAGATGTTGCTGCACCCATGGGCGCTGT 858
 Qy 281 AlaAlaGlyValMetHisLeuPheTyrValAlaGlyValSerArgAlaMetLeuLeu 300
 Db 859 GCAGCTGGCGTCAATGATCTCTTCTACGTCGGGCGGAGTGTCTGCTGCATGCTGCTG 918
 Qy 301 ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVal 320
 Db 919 CGCTCTTCTGCTGCATGAGAGGCGGAGGAGCATATGAGATGAAATGCCCTACTGCTGA 978
 Qy 321 TyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal 340
 Db 979 TATGTGCGCGGTGGCTTCCGCTTGAGCCCAAGAGATGGGAAAGTATGTTTGCAGTG 1038
 Qy 341 AspGlyLysLeuMetValSerGluAlaValGlnGlyValHisProAsnTyrPheTyr 360
 Db 1039 GATGGGGAATGATGATGTTACGAGGCGGTGCAGGCGCAAGTGCACCAACTACTTCTG 1098
 Qy 361 MetValSerGlyCysValGluProProSerTrpLysProGlnMetProProPro 380
 Db 1099 ATGGTCAGCGGTTCGTCGAGAGCCCGCCAGCTGAGAACCCAGCAGATGCCACGCCCA 1158
 Qy 381 GluGluProLeu 384
 Db 1159 GAAGAGCCCTTA 1170
 Db
 RESULT 7
 AAC84161
 ID AAC84161 standard; cDNA; 1205 BP.
 AC AAC84161:
 XX 19-MAR-2001 (first entry)
 DT
 XX
 DE Human sphingosine kinase encoding cDNA.
 XX
 KW Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic;
 KW antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
 KW human; ss.
 OS
 XX Homo sapiens.
 OS
 XX
 FT Key Location/Qualifiers
 FT CDS 33..1187
 FT /tag= a
 FT /product= "sphingosine kinase"
 XX
 PD WO200070028-A1.
 PD 23-NOV-2000.
 PE 12-MAY-2000; 2000MO-AU00457.
 XX
 PR 13-MAY-1999; 99AU-0000339.
 PR 08-JUL-1999; 99AU-0001504.
 XX
 PA (JOHN J) JOHNSON & JOHNSON RES PTY LTD.
 XX
 PI Plitson SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vadas MA;
 DR WPI; 2001-016227/02.
 DR P-PSDB; AAB48007.
 XX
 PT Novel sphingosine kinase protein and nucleic acid molecules for
 PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,

PT atherosclerosis, inflammation, meningitis, multiple sclerosis and
 PT septic shock -
 XX
 PS
 XX
 CC Claim 4; Fig 7a; 100pp; English.
 CC
 CC This cDNA encodes a human sphingosine kinase (SK) protein. The human SK
 CC protein, encoding nucleic acids and modulators are useful for modulating
 CC expression, functional activity or cellular functional activity of
 CC sphingosine kinase in a subject and also for treating a mammal by
 CC modulating the activity of SK. Diseases treated by regulating SK
 CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
 CC inflammation, meningitis, multiple sclerosis and septic shock.
 CC
 SQ Sequence 1205 BP; 196 A; 357 C; 400 G; 252 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,94e-188 Length: 1205
 Score: 1959.00 Matches: 374
 Percent Similarity: 98.18% Conservative: 3
 Best Local Similarity: 97.40% Mismatches: 7
 Query Match: 97.22% Indels: 0
 DB: Gaps: 0
 US-09-784-810A-2 (1-384) x AAC84161 (1-1205)
 Qy 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
 Db 33 ATGGATTCACAGCGGGGCGCCCGGGGGCTGTCCCGGGCCCTCCGCTGCTGCTGCTG 92
 Qy 21 LeuAsnProAlaArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
 Db 93 CTGAACCGCGCGCGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 152
 Qy 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluArgAlaHisAla 60
 Db 153 CTTTGGGCTGAGCGTGAATCTCTTCACTGATGCTCACTGAGGAGGAGGAGGAGGAGGAGG 212
 Qy 61 ArgGluLeuValArgSerGluLeuGluArgTyrPheAlaLeuValMetSerGly 80
 Db 213 CGGAGCTGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 272
 Qy 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTyrGluThrAla 100
 Db 273 GAGGCGCTGATGACAGAGTGTGAAGCGGCTCATGAGAGCGGCTGACTGGAGAGCGGCG 332
 Qy 101 IleGlnLysProLeuSerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu 120
 Db 333 ATCCAGAAAGCCCTGTGTAGCTCTCCAGAGGCTGTGGCAGACGCGCTGCACTTCTTG 392
 Qy 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
 Db 393 AACCATTTATGCTGGCTATGAGCGAGTGCACCAATGAAGACCTCTGACCAACTGCACGCTA 452
 Qy 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
 Db 453 TTGCTGTGCGCGCGGCTGTGCTACCACTGAACTCTGCTGTGACACGCGCTTCGGGG 512
 Qy 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu 180
 Db 513 CTGGCGCTTCTCTGTGCTGCTACAGCTGGCGCTGGGGCTTCATGCTGATGAGCACTAGAG 572
 Qy 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
 Db 573 AGTGAGAAGATACGCGCTGTGGGAGATCGGCTTCACTGTGGCACTTTCCTGGCGCTG 632
 Qy 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
 Db 633 GCAGCTTGGCGCACTTACCGGCGGCGGAGCTGCTTACTCTGTAGGAAGAGTGGCTTC 692
 Qy 221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro 240
 Db 693 AAGACACTGCTCCCGCTGTGTGTGCTCAGAGAGGCGCGGTATGATGACACTTGTGCCA 752


```

Db      541 AGTGAAGATGATCGGCTGCGGAGATGCCCTTCACTGGGACCTTCCTGCTGCTG
      |||.....
Oy      201 AAlaAlaLeuAlaArgThyTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe
      |||.....
Db      601 GCAGCCCTGGCGACCTACCGCGGCGGACCTGACCTCCCTGTAAGAAAGTGGTTC
      |||.....
Oy      221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro
      |||.....
Db      661 AAGACACCTGCTCCCGCTGGTGTTCAGACAGGCGCGGTGATGACACCTTGTGCCA
      |||.....
Oy      241 LeuGlnGlnGlnValProSerHisTrpGlnValValProAspGlnAspPheValLeuVal
      |||.....
Db      721 CTGGAGAGACCACTGCTCTTCACTGACAGAGTGTGCCGACAGAGACTTGTGCTAGTC
      |||.....
Oy      261 LeuAlaLeuLeuHisSerHisLeuAlaSerGlnMetPheAlaAlaProMetGlyArgCys
      |||.....
Db      781 CTGGCACTGCTGCACTGCACTGGGAGAGATGTTGCTGACCACTGGCCCTGT
      |||.....
Oy      281 AAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu
      |||.....
Db      841 GCAGCTGGCGTCATGATCTGTTTACGTGCGGCGGAGATGCTGCTGCACTGCTG
      |||.....
Oy      301 ArgLeuPheLeuAlaMetGlyLysGlyArgHisMetGlyTyrGlyCysProTyrLeuVal
      |||.....
Db      901 CCGCTCTTCTGGCCATGAGAGAGGAGGAGATGATGATGATGATGATGATGATGATG
      |||.....
Oy      321 TyrValProValValAlaPheArgLeuGlnProLysAspGlyLysGlyValPheAlaVal
      |||.....
Db      961 TATGTGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
      |||.....
Oy      341 AspGlyGlnLeuMetValSerGlnAlaValGlnGlnGlnValHisProSerTyrPheTyr
      |||.....
Db      1021 GATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
      |||.....
Oy      361 MetValSerGlyCysValGlnProProProSerTyrPlysProGlnGlnMetProPro
      |||.....
Db      1081 ATGCTACGCGGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
      |||.....
Oy      381 GlnGlnProLeu 384
      |||.....
Db      1141 GAAGAGCCCTTA 1152
      |||.....

```

RESULT 9

AAA50508 standard; cDNA; 1447 BP.

AAA50508:

05-DEC-2000 (first entry)

Human sphingosine kinase A cDNA.

Sphingosine kinase A; SKA; human; drug screening; infection;

antiinflammatory; antiallergic; anticancer; inflammation; allergy;

cancer; therapy; diagnosis; ds.

Homo sapiens.

Key Location/Qualifiers

FT CDS 7..1161

PN WO200052173-A2.

PD 08-SEP-2000.

PF 02-MAR-2000; 2000WO-CA00223.

PR 02-MAR-1999; 99US-0122516.

PA (ALX) NPS ALLELIX CORP.

PI Munroe D, Gupta A, Falzone GR;
 XX WPI: 2000-572185/53.
 DR P-PSDB; AAY06057.
 XX
 PT New human sphingosine kinase A, B and C polynucleotides and
 PT polypeptides useful in e.g. chromosome and gene mapping, and detecting
 PT inflammation or disease associated with abnormal levels of sphingosine
 PT kinase expression
 XX
 PS Disclosure: Fig 1; 81pp; English.

The present sequence is that of an isolated polynucleotide encoding human sphingosine kinase A (SKA, see AAY96057), an enzyme that phosphorylates sphingosine to form sphingosine 1-phosphate. The polynucleotide was isolated from an HeLa cDNA library by PCR amplification. The invention provides polynucleotides (see AAA50508-10) and polypeptides (see AAY6057-59) for the human sphingosine kinase (SK) homologues SKA, SKB and SKC. The polynucleotides may be used as hybridization probes, in the construction of PCR primers for chromosome and gene mapping, in the recombination of antisense DNA or RNA. They can be used to detect the expression of or to detect differences in gene sequence between normal and carrier or affected individuals. Host cells expressing SK can be used in drug screening. Human SK specific antibodies, inhibitors, ligands or their analogues can be used as bioactive agents to treat inflammation or disease including viral, bacterial or fungal infections, allergic responses, mechanical injury or associated with trauma, hereditary diseases, lymphoma or carcinoma, and other conditions which activate the genes of kidney, lung, heart, lymphoid or tissues of the nervous system.

Sequence 1447 BP; 256 A; 427 C; 465 G; 299 T; 0 other:

Alignment Scores:
 Pred. No.: 5, 2e-187 Length: 1447
 Score: 1946.00 Matches: 372
 Percent Similarity: 97.408 Conservative: 2
 Best Local Similarity: 96.888 Mismatches: 10
 Query Match: 96.588 Indels: 0
 DB: 21 Gaps: 0

US-09-784-810A-2 (1-384) x AAA50508 (1-1447)

```

Oy      1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu
      |||.....
Db      7 ATGATCCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG
      |||.....
Oy      21 LeuAsnProArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly
      |||.....
Db      67 CTGAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
      |||.....
Oy      41 LeuLeuAlaGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln
      |||.....
Db      127 CTTTGGCTGAGGCTGAATCTCTTCACTTACCTATGCTCAGTACGAGGAGAACCGCG
      |||.....
Oy      61 ArgGlnLeuValArgSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln
      |||.....
Db      187 CGGAGCTGCTGCGGCTGCGGAGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG
      |||.....
Oy      247 GACGGGCTGATGCAAGAGTGTGTAAGCGGCTCACTGAGCGGCTGACTGAGAGCGGC
      |||.....
Oy      101 TleGlnLysProLeuCysSerLeuProAlaGlySerGlyGlnAlaLeuAlaAlaSerLeu
      |||.....
Db      307 ATCCAGAAACCCCTGTATCTCCACAGAGCTCTGCAACGCGCTGGAGCTTCCTTG
      |||.....
Oy      121 AsnHisTyrAlaGlyTyrGlnGlnValThrAsnGlnAspLeuThrAsnCysThrLeu
      |||.....
Db      367 AACCATATGCTGCTATGAGCAGGTACCAATGAAAGCTCTGACCACTGACGAGCTA
      |||.....

```


PR 05-SEP-1997; 97US-0057627.
 PR 05-SEP-1997; 97US-0057634.
 PR 05-SEP-1997; 97US-0057643.
 PR 05-SEP-1997; 97US-0057646.
 PR 05-SEP-1997; 97US-0057649.
 PR 05-SEP-1997; 97US-0057654.
 PR 05-SEP-1997; 97US-0057656.
 PR 05-SEP-1997; 97US-0057660.
 PR 05-SEP-1997; 97US-0057663.
 PR 05-SEP-1997; 97US-0057763.
 PR 05-SEP-1997; 97US-0057769.
 PR 05-SEP-1997; 97US-0057774.
 PR 05-SEP-1997; 97US-0057777.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA;
 PI Fan P, Feng P, Ferlie AM, Fischer CL, Florence C;
 PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW;
 PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
 PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
 XX WPI: 1999-059865/05.
 DR P-PSDB; AAW88613.
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Claim 4: Page 343; 772pp; English.
 XX
 CC The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
 CC encoding human secreted proteins (AAW88534 to AAW88756). The secreted
 CC protein gene sequences are deposited with the ATCC under deposit numbers
 CC ATCC 97978, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
 CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
 CC cells comprising recombinant vectors containing the nucleic acid
 CC sequences are used for the recombinant production of the secreted
 CC proteins. The polynucleotide and amino acid sequences are useful for are
 CC useful for preventing, treating or ameliorating medical conditions e.g.
 CC by protein or gene therapy. Pathological conditions can be also
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the polynucleotides, based on
 CC which tissues they are most highly expressed in, and include developing
 CC products for the diagnosis or treatment of cancer, neurodegenerative
 CC disorders, developmental abnormalities and foetal deficiencies, blood
 CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
 CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
 CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
 CC osteoporosis, prostate diseases, obesity, disorders involving osteoclasts
 CC such as osteoporosis, arthritis or malignancies, diseases of testes,
 CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The present sequence represents a gene encoding a human secreted protein
 CC (see descriptor line for gene number and clone identification).
 CC
 XX
 SO Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other;
 Alignment Scores:
 Pred. No.: 4.72e-185 Length: 1533
 Score: 1927.00 Matches: 371
 Percent Similarity: 96.88% Conservative: 2
 Best Local Similarity: 96.36% Mismatches: 11
 Query Match: 95.63% Indels: 1
 DB: 20 Gaps: 0
 US-09-784-810a-2 (1-384) x AAV84490 (1-1533)
 OY 1 MetaspProAlaGlyLysProArgLysValLeuProArgProCysArgValLeuValLeu 20
 DB 73 ATGATTCACGAGGG 132
 OY 21 LeuAsnProArgGlyLysLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40

DB 133 CTGAACCGCGCGCGGCAAGGCAAGGCGCTTGCAGCTCCGAGAGTCAGCGAGCGCC 192
 OY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAlaGlnHisAla 60
 DB 193 CTTTGGGCTGAGGCTGAATCTCTTCACGCTGATGCTCAGCGAGCGCGGCAAGCGCG 252
 OY 61 ArgGluLeuValArgSerGluLeuGlnLeuArgTrpPheAlaLeuValMetSerGly 80
 DB 253 CGGAKCTGCTGCGGCTGAGAGAGCTGGGCGGCTGGAGCGCTTGTGGTATGTTGGA 312
 OY 81 AspGlyLeuMetHisGluValAlaValAsnGlyLeu-MetGluArgProAspTrpGlnTrpAl 100
 DB 313 GAGCGGCTGATGCAAGAGGCTGTAGAGCGGCTTATGAGAGCGGCTGAGCGAGCGCG 372
 OY 100 aileGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaAspLe 120
 DB 373 CATCCAGAGAGCCCTGTGTACCTCCAGAGCGCTGCGCAACGCGCTGCGAGCGCTTCT 432
 OY 120 uAsnHisTrpAlaGlyTrpGluGlnValThrAsnGluAspLeuThrPheAsnGlyThrLe 140
 DB 433 RAACCATTAATGCTGCTATGACAGGTCACCAATGAAGACCTCTGACCAACTGACAGCT 492
 OY 140 uLeuLeuCysArgProValLeuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160
 DB 493 ATTGCTGCGCGCGCGGCTGCTGCTACACCATGACCTGCTGCTGACAGCGCTTGGG 552
 OY 160 yLeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuG 180
 DB 553 GCTGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612
 OY 180 uSerAspTrpTrpArgTrpGluGlnGlyGluMetArgPheThrLeuGlyTrpPheAspLe 200
 DB 613 GAGTGAAGAGTATGCGGCTGCGGAGATGCGCTTCTGCTGCGACCTTCTGCGCT 672
 OY 200 uAlaAlaLeuArgThrTrpArgGlyArgLeuAlaThrLeuProValGlyArgValAla 220
 DB 673 GCGAGCGCTGCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 732
 OY 220 elyThrProAlaSerProValAlaValGlnGlnGlyProValAspAlaHisLeuValPr 240
 DB 733 CAAGCACCT 792
 OY 240 OLeuGlnGluGlnValProSerHisThrGlnValValProAspGluAspPheValLeuVa 260
 DB 793 ACTGAGAGAGGACGAGGCT 852
 OY 260 lLeuAlaLeuHisSerHisSerHisLeuAlaSerGluMetPheAlaIleProMetGlyArgCy 280
 DB 853 CCGGCACTGCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCTGCGACCT 912
 OY 280 sAlaAlaGlyValMetHisLeuPheTrpValAlaArgAlaGlyValSerAlaGluMetLeu 300
 DB 913 TGCAGCTGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
 OY 300 uArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTrpGluCysProTrpLeuVa 320
 DB 973 GCGCCTCTCTCTGCGCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1032
 OY 320 lTrpValProValAlaAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVa 340
 DB 1033 ATATGTCGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1092
 OY 340 lAspGlyLysLeuMetValSerGluAlaValGlnGlyLysGlnValHisProAsnTrpPheTr 360
 DB 1093 GGATGGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1152
 OY 360 pMetValSerGlyCysValGluProProPheSerTrpLysProGlnGlnMetProProPr 380
 DB 1153 GATGTCAGCGGCTGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1212
 OY 380 oGluGluProLeu 384
 |||||

Db 1213 AGAGAGCCCTTA 1225

RESULT 12
ABAB3273
ID ABAB3273 standard; cDNA; 1533 BP.
XX
AC ABAB3273;
XX
DT 07-FEB-2002 (first entry)
XX
DE Human secreted protein gene 80 SEQ ID NO:90.
XX
KW Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
KW dermatological; immunosuppressive; anti-inflammatory; immunostimulant;
KW cytostatic; cardiant; vascular; anti-angiogenic; ophthalmological;
KW neuroprotective; neurotropic; anticonvulsant; antialzheimer's; vulnary;
KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW Gaucher's disease; cardiovascular disease; Sclimlar syndrome; chemotaxis;
KW Chaga's cardiomyopathy; coronary arteriosclerosis; angiotensin II disorder;
KW corneal graft neovascularisation; diabetic retinopathy; regeneration;
KW neurological disorder; Huntington's chorea; Alzheimer's disease;
KW Parkinson's disease; infectious disease; chromosome 17; ss.
OS Homo sapiens.
XX
PN MO200162891-AZ.
XX
PD 30-AUG-2001.
XX
PE 21-FEB-2001; 2001WO-US05614.
XX
PR 24-FEB-2000; 2000US-184836P.
XX
PR 29-MAR-2000; 2000US-193170P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PA NI J, Ebner R, Lafleur DM, Moore PA, Olsen HS, Rosen CA;
PI Ribben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferris AM, Fan P;
PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;
PI Zeng Z, Greene JM;
XX
DR WPI; 2001-625724/72.
XX
DR P-PSBB; ABAB50380.
XX
XX Nucleic acids encoding 207 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating, e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
PT
XX
PS Claim 1; Page 965; 1533pp; English.
XX
XX ABAB50301 to ABAB51287 and ABAB3194 to ABAB3441 represent human secreted
CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
CC activities based on the tissues and cells the genes are expressed in.
CC Example of these activities include: immunomodulatory; antisclerotic;
CC dermatological; immunosuppressive; anti-inflammatory; immunostimulant;
CC anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;
CC neuroprotective; neurotropic; anticonvulsant; antialzheimer's; vascular;
CC antiparkinsonian; antimicrobial; and vulnary. (I) and (II) can be used
CC in gene therapy and vaccine production. (I) and (II) can be used in the
CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC Gaucher's disease), cardiovascular diseases (e.g. Sclimlar syndrome,
CC Chaga's cardiomyopathy, and coronary arteriosclerosis), angiotensin
CC disorders (e.g. corneal graft neovascularisation and diabetic
CC retinopathy), neurological disorders (e.g. Huntington's chorea,
CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC for promoting wound healing, regeneration and/or chemotaxis. ABAB3185 to
CC ABAB3193 and ABAB50300 represent sequences used in the exemplification of
CC the present invention.

XX
SQ Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other;
XX
Alignment Scores:
Pred. No.: 4,72e-185 Length: 1533
Score: 1927.00 Matches: 371
Percent Similarity: 96.88% Conservative: 2
Best Local Similarity: 96.36% Mismatches: 11
Query Match: 95.63% Indels: 1
DB: 22 Gaps: 0

US-09-784-810a-2 (1-384) x ABAB3273 (1-1533)

1 MetaspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
73 ATGATCCAGCGGGGCGGGCCCGGGGGCTCCCGCGCGCCCTGCTGCTGCTGCTGCTG 132
21 LeuAnpProArgGlyGlyValLeuValLeuValLeuValLeuValLeuValLeuValLeu 40
133 CTGAACCG 192
41 LeuValAlaGlyAlaGlyLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60
193 CTTTGGCTGAGCTGGAATCTCTTACGCTGATGCTCACTGACGCGCGGAAACCAACGCG 252
61 ArgGluLeuValArgSerGluGluLeuGlyArgGlyArgGlyArgGlyArgGlyArgGly 80
253 CCGGARGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
81 AspGlyLeuMetHisGlyValValAlaAsnGlyLeu-MetGluArgProAspTrpGluTrpAl 100
313 GAGGGCTGATCAGCAGAGTGTGTGAGAGGGCTTATGAGAGGGCTGATGAGAGAGAGAG 372
100 alleGlnLysProLeuGlySerLeuProAlaGlySerLysAlaLeuAlaLeuAlaLeu 120
373 CATCCAG 432
120 uAnHisTyrAlaGlyTyrGluGluValThrAsnGluAspLeuThrAsnGlyThrLe 140
433 RAACCATTAATGCTGCTATRACAGAGTCAACCAATGAGACCTCTGACCACTGACAGCT 492
140 uLeuLeuGlyArgProValLeuSerProMetAsnLeuLeuSerLeuHisTyrAlaSerG 160
493 ATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
160 yLeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeu 180
553 GCTGCGCT 612
180 uSerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArg 200
613 GAGTGAAGAATACGGGGCTGCGGAGATGCGCTTCACTGCGGACCTTCTCTCTCTCTCT 672
200 uAlaAlaLeuArgThrTyrArgGlyValArgLeuAlaTrpLeuProValGlyArgValGly 220
673 GCGCGCCCTGCGCACCTACCGCGCGCGAGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 732
220 eLysThrProAlaSerProValValAlaGlnGlnGlyProValAlaHisLeuValPr 240
733 CAAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792
240 oLeuGluGluGluValProSerHisTyrPheValValAlaProAspGluAspPheValLeu 260
793 ACTGAG 852
260 lLeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgG 280
853 CTTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912
280 sAlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeu 300
913 TCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972

OY 300 uArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuVa 320
 |||||||
 Db 973 GGGCTCTTCCCTGGCCATGAGAGAGGCGCATGTGAGTGAATGCCCTACTTGCT 1032
 OY 320 lTyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVa 340
 |||||||
 Db 1033 ATATGTGCCCGGTGTCGCTTCCGCTTGAGAGCCCAAGATGGAAAGGTGTGTGCACT 1092
 OY 340 lAspGlyLeuLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTr 360
 |||||||
 Db 1093 GGATGGGGAATGATGATGATGAGAGCGCCGTCAGGCGCCAGTGCACCCAACTACTCTCG 1152
 OY 360 pMetValSerGlyCysValGluProProProSerTrpLysProGlnGlnMetProProPr 380
 |||||||
 Db 1153 GATGCTCAGCGGTTGGCTGGAGCCCCCGCCAGCTGGAAAGCCCAAGATGCCACCGCC 1212
 OY 380 oGluGluProLeu 384
 |||||||
 Db 1213 AGAAGAGCCCTTA 1225
 RESULT 13
 AAD14425
 ID AAD14425 standard; cDNA; 1759 BP.
 AC AAD14425;
 XX 01-NOV-2001 (first entry)
 DE Mouse consensus sphingosine kinase (Sphk) cDNA.
 XX
 KW Mouse; sphingosine kinase; Sphk; restenosis; ischaemia; gene therapy;
 KM antisenase therapy; cancer; sphingolipid; signalling molecule; apoptosis;
 KW cytosolic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
 XX leukaemia; vasotropic; cell proliferative disorder; vascular disease; ss.
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 327..1475
 FT /tag= a
 FT /product= "Mouse sphingosine kinase (Sphk) protein"
 XX
 PN WO200160990-A2.
 PD 23-AUG-2001.
 XX
 PF 14-FEB-2001; 2001WO-US04789.
 XX
 PR 14-FEB-2000; 2000US-0182360.
 PR 22-MAR-2000; 2000US-0191261.
 XX
 PA (CURA-) CURAGEN CORP.
 PA (GETH) GENENTECH INC.
 XX
 PI Rastelli L;
 DR MPI: 2001-514770/56.
 DR P-PSDB: AAE07883.
 XX
 PT An isolated sphingosine kinase polypeptide useful for treating a
 PT Sphk-associated disorder especially cancer, restenosis or ischemia in a
 PT human -
 XX
 PS Example 1; Page 92; 107pp; English.
 XX
 CC The present invention relates to sphingosine kinase (Sphk) polypeptides
 CC and nucleic acids encoding them. Sphk is useful for treating a Sphk-
 CC associated disorder especially cancers such as leukaemia, lymphoma,
 CC ovarian, breast, lung, colon, testicular, stomach and skin,
 CC atherosclerosis, restenosis or ischaemia and cell proliferative disease
 CC or disorder associated with vascular diseases. Sphk gene is used in gene
 CC therapy and antisenase-therapy. Sphingolipids serving as signalling
 CC molecules, have recently emerged as regulators of cell growth.

CC differentiation, diverse cell phenotypes and cell death. Activation of
 CC Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
 CC endothelial cells. The present sequence is consensus mouse sphingosine
 CC kinase (Sphk) cDNA.
 XX
 SQ Sequence 1759 BP; 357 A; 495 C; 523 G; 384 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,566-151 Length: 1759
 Score: 1594.50 Matches: 303
 Percent Similarity: 88.25% Conservative: 35
 Best Local Similarity: 79.11% Mismatches: 44
 Query Match: 79.13% Indels: 1
 DB: 22 Gaps: 1
 US-09-784-810A-2 (1-384) x AAD14425 (1-1759)
 OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
 |||||||
 Db 327 ATGGAACCAAGTGAATGCCCTGAGAGACGCTCCACAGCCATGCAAGTGTGCTG 386
 OY 21 LeuAsnProArgGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
 |||||||
 Db 387 CTGAACCCCAAGGCTGGCAAGGCAAGGCTCTGACAGCTTCCAGAGCCGTGTGACAGCC 446
 OY 41 LeuLeuAlaGluAlaGluLeuSerPheTrpLeuMetLeuThrGluArgAsnHisAla 60
 |||||||
 Db 447 TTCCTGGAGAGGAGATTAACCTTTAACTGATATCTACCAAGGAAAGCAATGCG 506
 OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
 |||||||
 Db 507 AGGAGCTGTGTGTGACAGAGAGTGGGTCACTGGAGAGCCCTGGACATGATGTCGGT 566
 OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
 |||||||
 Db 567 GATGCTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 626
 OY 101 lIleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaLeuAlaSerLeu 120
 |||||||
 Db 627 ATCCAGAAACCCCTGTGTGCTGCTCCCTGGAGGCTCCGCAATGCCCTGACCTTCTGTG 686
 OY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnGlyThrLeu 140
 |||||||
 Db 687 AACCACTATGCTGGGTGACAGCAGCAGTGAATGAAGACCTGCTCATCACTGACACAG 746
 OY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
 |||||||
 Db 747 CTGTTGTCGCCGCCGCCCTGTACACCATGAACTGTGCTCCCTGACACTGCTCTGGG 806
 OY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAlaLeuGlu 180
 |||||||
 Db 807 CTGCGGCTATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
 OY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyTyrPheLeuArgLeu 200
 |||||||
 Db 867 AGTGAAGATACAGGCGCTTGGGGAGATGCTTTCACAGTGCGGACCTCTTGGCCGA 926
 OY 201 AlaAlaLeuArgTrpTyrArgGlyArgLeuAlaTrpLeuProValGlyArgValGlyPhe 220
 |||||||
 Db 927 GCAAGCTCTGCAATCACCAGCCCAAGGCGCTTCACTTCTGTGGAAGACTGTGCGCTCT 986
 OY 221 LysThrProAlaSerProValValAlaGlnGlnGlnGlyProValAspAlaHisLeuValPro 240
 |||||||
 Db 987 AAGAGACCCGCTCT---ACACTGGGCGAAAGGCGCCGCTGGACACACACTGTGTCTCT 1043
 OY 241 LeuGluGluGluValProSerHisTrpGlnValValProAspGluAspPheValLeuVal 260
 |||||||
 Db 1044 CTGGAGGAGCAGAGTGCTTCTCATTTGAGTGTGTGATACAGAACAGAGACTGTCTCGTGG 1103
 OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaPrometGlyArgCys 280
 |||||||
 Db 1104 CTGGCTGTCTACACACACCACTGAGCTCGAGCTGTTCGACAGCACTGAGGCGCTGT 1163

		CC	activity or expression are used: (i) to reduce cell proliferation,
		CC	specifically for treating cancer, and (ii) to treat diseases associated
		CC	with abnormal cell migration or motility, particularly cancer, restenosis
		CC	or diabetic neuropathy (but also atherosclerosis, stroke and Alzheimer's
		CC	disease), whereas agents that stimulate sphingosine kinase can be used
		CC	to treat conditions associated with reduced cell proliferation, e.g.
		CC	developmental retardation.
XX	Sequence 1559 BP; 333 A; 420 C; 447 G; 359 T; 0 other;		
SQ			
	Alignment Scores:		
	Pred. NO.:	1.76e-150	Length: 1559
	Score:	1585.50	Matches: 301
	Percent Similarity:	88.86%	Conservative: 34
	Best Local Similarity:	79.84%	Mismatches: 41
	Query Match:	78.68%	Indels: 1
	DB:	21	Gaps: 1
US-09-784-B10A-2 (1-384)	x AAZ47168 (1-1559)		
OY	7 ProArGlyValLeuProArgProcysArgValLeuValLeuLeuAsnProArgglyLy	26	
Db	118 CTCGAGACGGCTGCCACCGGCCTATGCAAGTGTGGTGCTGTGAACCCCACGGGTGC	177	::::
OY	27 LysGLyLysAlaLeuGlnLeuPheArgSerHisValGlnProLeuLeuLaGluLaGlu	46	
Db	178 AAGGCGAAAGCCTTCGACTCTTCACAGCGCGGTGCAGCCCTTCGAGAGGGCAGAG	237	
OY	47 IleSerPheThrLeuMetLeuThrGlnArgArgnsHisAlaTrgIleuValAcySer	66	
Db	238 ATAACTTTAACTGATTACTACCAGAACGGAGAACAATGCGAGAGACTGGTGTA	297	
OY	67 GluGluLeuGlyArgTrpAspAlaLeuValAlaMetSerGlyAspGlyLeuMethIsglu	86	
Db	298 GAGAGATTGGGTCACTGGAGAGCCCTGCAGCATGTCGGTGATGCTGATGCATGAA	357	
OY	87 ValValAsnGlyLeuMetLeuLuarGrpProAspTrpGluTrfAlaIleGlnLysProLeuCys	106	
Db	358 GTGTGAATGGCGTTAATWGAGACGGCACAGCTGGAGACTGGCCATCCAGAACCCCTGT	417	
OY	107 SerLeuProAlaGlySerGlyAsnAlaLeuAlaaserLeuAsnHisTyraIagLYTYR	126	
Db	418 AGCCTCCCTGGAGSGCTCCGGCAATGGCGTAGAGCGTTCTGTAAACCATATGCTGGGTAC	477	
OY	127 GluGlnValThrAsnGlnLysPheLeuThrAnscystHrLeuLeuLeuCysArgProval	146	
Db	478 GAGAGGTGACTAATGAGACGTGCTCATCACTCACTGACACTGCTGTGTGGCCGCGGC	537	
OY	147 LeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerVal	166	
Db	538 CTGTACCCCANAGACCTGCTGTCCTGCACACAGCTTTGGGGTGCGGCTCATTTCTGTG	597	
OY	167 LeuSerLeuAlaTrpGlyPheIleAlaaspValAspLeuGluserrAsplytyrArgar	186	
Db	598 CTCACTGTGCTCGGGGCTTTGTGTCAGAGTGGACTCGAGAGTGAAGAATGACAGGGCG	657	
OY	187 LeuGlyGluMetArgPheThrLeuGlyTYH-PheLeuArgLeuAlaIleuAlaArgTYrryr	206	
Db	658 TTGGGGGGAATTCGTTTCACATGGGACACTTTCTTGCTGTACAGACCTGGCCAGCTTAC	717	
OY	207 ArgGlyArgLeuAlaThrLeuProValGlyVargValGlyPheUsthrProalaSerPro	226	
Db	718 CAAGCGCAAATGCGCTTACCTTCCTGTAGAACCTTGGCSCTTAAGAAGCCCCGCTCT--	774	
OY	227 ValValAlaGlnGlnGlyProValAspAlaHisLeuValProleuglmgInuValPro	246	
Db	775 ACACGTGTGCAGAGGGCCCGCTGCACACACACACCTTTGCTCTGGAGAGCCAGTGCCT	834	
OY	247 SerHisTrpGlnValValProAspGlnSppHeValLeuValLeuAlaLeuHisser	266	
Db	835 TTCATATTGACTGTGTGTACAGAACAGACTTCGTCCTGGTGGCTGTGTCTACACACC	894	
OY	267 HisLeuAlaSerGluMetPheAlaAlaProMetGlyArgcylAlaIlaGlyValMetHis	286	


```

Db      1163 CACCTGAGCTCCGAGCTGTTTGCAGACCCATGGGCCCTGTGAAGCTGTATGCAT 1222
QY      287 LeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPheLeuAlaMet 306
        |||||||
Db      1223 CTGTCTACGTACGTGCGGGGGTGTCAAGGGCTGCGCTGCGCTCTCTCTGCGCATG 1282
QY      307 GluLysGlyArgHisMetGluTyrGluCysProTyrLeuValTyrValProValValAla 326
        ::|||
Db      1283 CAGAGGGCCAGCATATGGAAGTGTGACCTCCATACCTGTTTCATGTGCGGTGCT 1342
QY      327 PheArgLeuGluProLysAspGlyLysGlyValPheAlaValAspGlyGluLeuMetVal 346
        |||||||
Db      1343 TTCCGCTGTGAGCCAGAGGAGGCGGTGTTTCTGTGTGATGAGAGCTGATGTA 1402
QY      347 SerGluAlaValGlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysVal 366
        |||||||
Db      1403 TGTGAAGCTGTGAGGCGCAAGTGCAACCACTTGTGATGGTGTGTGGCAGACAGA 1462
QY      367 GluProProProSerTrpLysProGlnGlnMetProProProGluGluPro 383
        ::|||
Db      1463 GATGCCCATCCGGCGGAGCTCCCGGGGGGCCACCTCCAGAGACCA 1513

```

Search completed: July 12, 2003, 18:17:50
 Job time : 309 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2003, 18:09:03 ; Search time 3410 Seconds

(without alignments)
3277.266 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGRGVLPKPCRVLYL.....CVEPPPSKPKQMPPEEPL 384

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODE=frame_plus_p2n.model -DEV=xlh
-O=/cgn2_1/USPPO/US09784810/runat_11072003_103441_26557/app-query.fasta.1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=txt -NORM=ext -HEAPSIZE=500 -MITEM=0 -MAXLEN=2000000000
-USER=US09784810.ecgn.1.1.2456-etunat.11072003_103441_26557 -ICPU=3
-NO_MAP -LARGEUSER -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vl.*
30: em.htg.hum.*
31: em.htg.in.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2015	100.0	1600	6	AX224379	AX224379 Sequence
2	1962	97.4	1155	6	AX127642	AX127642 Sequence
3	1962	97.4	1693	9	AF266756	AF266756 Homo sapi
4	1962	97.4	1719	6	AX127641	AX127641 Sequence
5	1962	97.4	1772	9	AK023393	AK023393 Homo sapi
6	1962	97.4	1821	9	AK022402	AK022402 Homo sapi
7	1959	97.2	1173	9	AF200328	AF200328 Homo sapi
8	1955	97.0	1824	9	BC008040	BC008040 Homo sapi
9	1954	97.0	1783	9	AF238083	AF238083 Homo sapi
10	1945	96.5	1869	9	BC009419	BC009419 Homo sapi
11	1944	96.5	2502	9	AK095578	AK095578 Homo sapi
12	1880.5	93.3	1428	9	HS2445504	AJ245504 Homo sapi
13	1813	90.0	183443	2	AC068145	AC068145 Homo sapi
14	1789.5	88.8	2015	9	AB046025	AB046025 Macaca fa
15	1594.5	79.1	1759	6	AX224381	AX224381 Sequence
16	1587.5	78.8	209861	2	AC021196	AC021196 Homo sapi
17	1585.5	78.7	1146	10	AF415213	AF415213 Mus muscu
18	1585.5	78.7	1559	6	AX287138	AX287138 Sequence
19	1585.5	78.7	1559	6	AF068749	AF068749 Mus muscu
20	1582.5	78.5	1815	6	AX287137	AX287137 Sequence
21	1582.5	78.5	1815	10	AF068748	AF068748 Mus muscu
22	1555.5	77.2	1842	10	AB049573	AB049573 Rattus no
23	1555.5	77.2	1895	10	AB049572	AB049572 Rattus no
24	1555.5	77.2	1943	10	AB049571	AB049571 Rattus no
25	1555.5	77.2	2501	10	AB049574	AB049574 Rattus no
26	1555.5	77.2	2648	10	AB049575	AB049575 Rattus no
27	1454	72.2	1192	9	BC004112	BC004112 Homo sapi
28	1451.5	72.0	110000	2	AL606505	AL606505 Mus muscu
29	1451.5	72.0	112037	10	AL645851	AL645851 Mouse DNA
30	875	43.4	2698	10	AF245448	AF245448 Mus muscu
31	874	43.4	1854	10	AF415214	AF415214 Mus muscu
32	874	43.4	2416	10	BC006641	BC006641 Mus muscu
33	856.5	42.5	2380	9	AF245447	AF245447 Homo sapi
34	856.5	42.5	2731	9	BC010671	BC010671 Homo sapi
35	856.5	42.5	2875	6	AX086301	AX086301 Sequence
36	856.5	42.5	2875	9	HS0801669	AL136701 Homo sapi
37	856.5	42.5	3012	9	BC006161	BC006161 Homo sapi
38	731	36.3	200146	2	AC073774	AC073774 Mus muscu
39	685	34.0	119631	9	AC008888	AC008888 Homo sapi
40	685	34.0	179372	2	AC040922	AC040922 Homo sapi
41	673.5	33.4	38149	9	AC022154	AC022154 Homo sapi
42	609.5	30.2	254729	2	AC122611	AC122611 Rattus no
43	550	27.3	2629	3	AY119652	AY119652 Drosophill
44	550	27.3	46594	2	AC013897	AC013897 Drosophill
45	550	27.3	164415	3	AC023718	AC023718 Drosophill

RESULT 1

ALIGNMENTS

AX224379
LOCUS AX224379 1600 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0160990.
ACCESSION AX224379
VERSION AX224379.1 GI:15554631
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Rastelli, L.
TITLE Novel sphingosine kinases
JOURNAL Patent: WO 0160990-A 1 23-AUG-2001;
Curagen Corporation (US); GENENTECH, INC. (US)
FEATURES
source 1..1600
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 265 a 492 c 531 g 311 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 2,37e-168 Length: 1600
Score: 2015.00 Matches: 384
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-784-810A-2 (1-384) x AX224379 (1-1600)
QY 1 MetaspProAlaGlyGlyProAlaGlyValLeuProAlaGlyProCysArgValLeuValLeu 20
DB 176 ATGGATCAAGCGGG 235
QY 21 LeuAsnProAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 40
DB 236 CTGAACCCCG 295
QY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
DB 236 CTTTGGCTGAGGCTGAATCTCTTCACGGCTGAGTGCCTGAGTGCCTGAGTGCCTGAGTGCCT 355
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValMetSerGly 80
DB 356 CGGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 415
QY 81 AspGlyLeuMetHisGlyValValAsnGlyLeuMetGluArgProAspTrpGluTrpAla 100
DB 416 GACGGGCTGAGTGCAGCGAGGTGGTGAACGGGCTGAGTGCAGCGGCTGAGTGCAGCGGCTG 475
QY 101 IleGlnIysProLeuGlySerLeuProAlaGlySerGlyAsnAlaLeuAlaSerLeu 120
DB 476 ATCCAGAGAGCCCTGCTGAGCTCCAGAGGCTCTGGCAACGGGCTGAGGCTCTGCTG 535
QY 121 AsnHisTrpAlaGlyTrpGluGlnValThrAsnGluAspLeuLeuThrAsnGlyTrpLeu 140
DB 536 AACCATTTATGCTGCTATGAGCAGGTGCACCAATGAAGACCTCTGACCAACCTGACGCTA 595
QY 141 LeuLeuAspArgProValLeuSerProMetAsnLeuLeuSerLeuHisTrpAlaSerGly 160
DB 596 TTGCTGTCGCGCGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 655
QY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu 180
DB 656 CTGGCGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
QY 181 SerAspLysTrpArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 200
DB 716 AGTATAGATGATGCGGCTGCGGGGAGATGCGCTTCACTGCGGACACTTCTGCGCTGCG 775
QY 201 AlaAlaLeuArgTrpTrpArgGlyArgGlyLeuAlaTrpLeuProValGlyArgValGlyPhe 220

DB 776 GCACCCCTGGCGACACCTACCGCGCGGAGTGCCTGCTGCTGAGTGCCTGCTGCTGCTGCTGCT 835
QY 221 LysThrProAlaSerProValValAlaGlnGlnGlyProValAspAlaHisLeuValPro 240
DB 836 AAGACACCTGCTTCCCGGCTTGTGGTCCAGCGAGGGCGGCTGATGACACCTGCTGCTGCA 895
QY 241 LeuGlnGlnGlnValProSerHisTrpGlnValValProAspGluAspPheValLeuVal 260
DB 896 CTGAGAGAGAGAGTGCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
QY 261 LeuAlaLeuLeuHisSerHisSerHisLeuAlaSerGluMetPheAlaAlaPrometGlyArgCys 280
DB 956 CTGGACATGCTGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 1015
QY 281 AlaAlaGlyValMetHisLeuPheTrpValArgAlaGlyValSerArgAlaMetLeuLeu 300
DB 1016 GCACCTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1075
QY 301 ArgLeuPheLeuAlaMetGluGlyArgHisMetGluTrpGlyGlyProTrpTrpVal 320
DB 1076 GCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
QY 321 TyrValProValValAlaPheArgLeuGluProLysAspGlyGlyValPheAlaVal 340
DB 1136 TATGTGCCCTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1195
QY 341 AspGlyGluLeuMetValSerGluAlaValGlnGlyGlnValHisProAsnTrpPheTrp 360
DB 1196 GATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1255
QY 361 MetValSerGlyCysValGluProProProSerTrpLysProGlnGlnMetProProPro 380
DB 1256 ATGGTTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1315
QY 381 GlnGluProLeu 384
DB 1316 GAAAGCCCTT 1327
RESULT 2
LOCUS AX127642 1155 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 2 from Patent WO0131029.
ACCESSION AX127642
VERSION AX127642.1 GI:14134308
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Allen, J., Gosink, M., Melendez, A.J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 2 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source 1..1155
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 185 a 349 c 382 g 239 t
ORIGIN
Alignment Scores:
Pred. No.: 7.44e-164 Length: 1155
Score: 1962.00 Matches: 375
Percent Similarity: 98.18% Conservative: 2
Best Local Similarity: 97.66% Mismatches: 7
Query Match: 97.37% Indels: 0
DB: Gaps: 0
US-09-784-810A-2 (1-384) x AX127642 (1-1155)
QY 1 MetaspProAlaGlyGlyProAlaGlyValLeuProAlaGlyProCysArgValLeuValLeu 20

DB 390 CTTTGGCTGAGGCTGAAATCTCTTCACGCTGATGCTCACTAGCGGGGAGAACACGCG 449
OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValValMetSergly 80
DB 450 CGGAGCTGGTGGCTCGAGAGACTGGCGCGGAGCCCTGCTGCTGCTGCTGCTGA 509
OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpAla 100
DB 510 GAGGGCTGATGACAGAGAGTGTACAGGGGCTATGGAGGGGCTGACTGAGGAGACCGCC 569
OY 101 IllegInlyProLeuCySerLeuProAlaGlySerglyAsnAlaLeuValAsnSerleu 120
DB 570 ATCCAGAACCCCTGTGTATGCTCCAGACAGCTCTGGACCGCTGGACACTTCTCTTG 629
OY 121 AsnHisTyrAlaGlyTyrGluGluValThrAsnGluAspLeuThrAsnCystrleu 140
DB 630 AACCATATAGCTGGCTATGAGACAGTACACAGTAAGAACCTCTGACCAACAGACGCTA 689
OY 141 LeuLeuCyArgProValLeuSerProMetAsnLeuSerLeuHisThrAlaSergly 160
DB 690 TTGCTGGCGCGGCTGCTGCTGACCCATGACCTGCTCTGTCACACGCGCTCGGGG 749
OY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspAlaSerleu 180
DB 750 CTGCGCTCTTCTGCTGCTGACGCTGGCGCTGGGGCTTCATGTCTGATGTGACCTAGAG 809
OY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgleu 200
DB 810 AGTGAAGATATCGCGCTGCTGGGGAATGCGCTCACTGTGGCCCTTCCTGGGTGTG 869
OY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValAla 220
DB 870 GCAGCCTCGACCACTACCGCGCGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 929
OY 221 LysThrProAlaSerProValValGlyGlyGlyProValAspAlaHisLeuValPro 240
DB 930 AGACACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 989
OY 241 LeuGluGluGluValProSerHisThrPheValValProAspGluAspPheValLeuVal 260
DB 990 CTGGAGAGACCGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1049
OY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGlyMetPheAlaAlaProMetGlyArgGly 280
DB 1050 CTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1109
OY 281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeu 300
DB 1110 GCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1169
OY 301 ArgLeuPheLeuAlaMetGluLysGlyArgHisMetGluTyrGlyLysProTyrLeuVal 320
DB 1170 CGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1229
OY 321 TyrValProValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal 340
DB 1230 TAGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1289
OY 341 AspGlyGluLeuMetValSerGluAlaValGlyGlyGluValHisProAsnTyrPheTrp 360
DB 1290 GATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349
OY 361 MetValSerGlyCyValGluProProSerTrpLysProGluGluMetProProPro 380
DB 1350 ATGGTACGCGGTGGTGGAGACCCCGCCAGCTGAGAGCCAGAGATGACAGCCGCA 1409
OY 381 GluGluProLeu 384
DB 1410 GAAAGACCTTA 1421

LOCUS AX127641 1719 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0131029.
ACCESSION AX127641
VERSION AX127641.1 GI:14134307
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Allen, J., Gosink, M., Melendez, A.J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 1 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source location/Qualifiers
1..1719
BASE COUNT 305 a 529 c 556 g 329 t
ORIGIN
Alignment Scores:
Pred. No.: 1,23e-163 Length: 1719
Score: 1962.00 Matches: 375
Percent Similarity: 98.18% Conservative: 2
Best Local Similarity: 97.66% Mismatches: 7
Query Match: 97.37% Indels: 0
Gaps: 0
US-09-784-810A-2 (1-384) x AX127641 (1-1719)
OY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCyArgValLeuValLeu 20
DB 270 ATGATATCAGCG 329
OY 21 LeuAsnProArgGlyGlyGlyGlyGlyAlaLeuGluLeuPheArgSerHisValGluPro 40
DB 330 CTGAACCCG 389
OY 41 LeuLeuAlaGluAlaGluLeuSerPheThrLeuMetLeuThrGluAlaArgArgAsnHisAla 60
DB 390 CTTTGGCTGAGGCTGAAATCTCTTCACGCTGATGCTCACTAGCGGGGAGAACACGCG 449
OY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTrpAspAlaLeuValValMetSergly 80
DB 450 CGGAGCTGGTGGCTCGAGAGACTGGCGCGGAGCCCTGCTGCTGCTGCTGCTGA 509
OY 81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpAla 100
DB 510 GAGGGCTGATGACAGAGTGTACAGGGGCTATGGAGGGGCTGACTGAGGAGACCGCC 569
OY 101 IllegInlyProLeuCySerLeuProAlaGlySerglyAsnAlaLeuValAsnSerleu 120
DB 570 ATCCAGAACCCCTGTGTATGCTCCAGACAGCTCTGGACCGCTGGACACTTCTCTTG 629
OY 121 AsnHisTyrAlaGlyTyrGluGluValThrAsnGluAspLeuThrAsnCystrleu 140
DB 630 AACCATATAGCTGGCTATGAGACAGTACACAGTAAGAACCTCTGACCAACAGACGCTA 689
OY 141 LeuLeuCyArgProValLeuSerProMetAsnLeuSerLeuHisThrAlaSergly 160
DB 690 TTGCTGGCGCGGCTGCTGCTGACCCATGACCTGCTCTGTCACACGCGCTCGGGG 749
OY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspAlaSerleu 180
DB 750 CTGCGCTCTTCTGCTGCTGACGCTGGCGCTGGGGCTTCATGTGCTGATGAGACCTAGAG 809
OY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgleu 200
DB 810 AGTGAAGATATCGCGCTGCTGGGGAATGCGCTCACTGTGGCCCTTCCTGGGTGTG 869
OY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValAla 220

Db 870 GCACCCCTGCGACCTACCGCGCGGAGCTGCTACCTCCCTGAGAGAGTGGTTCC 929

Qy 221 LysThrProAlaSerProValValAlaGlnGlnGlyProValAlaPheAlaHisLeuValPro 240

Db 930 AAGACACCTGCTCCCGCGTGTGTGTCACAGAGGCGCGGATGACACCTTGTGCA 989

Qy 241 LeuAlaGluGlnValProSerHisTropInValValProAspGluAspPheValLeuVal 260

Db 990 CTGAGAGAGCCAGTGCCTCTCAGTGCAGACAGTGTGCTCCGCGAGAGCTTGTCTAGTC 1049

Qy 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaPrometGlyArgCys 280

Db 1050 CTGCACCTGCTGACCTGCGACCTGAGATGATGTTGCTGCACCCATGCGCGCTGT 1109

Qy 281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu 300

Db 1110 GCACCTGCGTCATGATCTCTTACGCTGCGGCGGAGTGTCTCGTCCATGCTGTG 1169

Qy 301 ArgLeuPheLeuAlaMetGluLysGlyValArgHisMetGluTyrGluCysProTyrLeuVal 320

Db 1170 CGCCTCTTCCCTGCGCATGAGAGGCGAGCATATGAGTATGATGCCCTACTTGTGTA 1229

Qy 321 TyrValProValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaVal 340

Db 1230 TATGTGCCCGTGTGCTGCGCTTCCGCTGAGCCCAAGATGGGAAAGTGTGTGCACTG 1289

Qy 341 AspGlyGluMetValSerGluAlaValGlnGlyGlnValHisProAsnTyrPheTyrP 360

Db 1290 GATGGGGAATGATGATTAGCGAGGCGCTGACGCGCCAGTGCACCCAACTACTCTCGG 1349

Qy 361 MetValSerGlyCysValGluProProProSerTyrPyrGlnGlnMetProProPro 380

Db 1350 ATGGTCACCGGTTGCTGCTGAGACCCCGCCAGCTGGAAGCCCGACAGATGCCACGCCA 1409

Qy 381 GluGluProLeu 384

Db 1410 GAAGAGCCCTTA 1421

RESULT 5

AK023393

LOCUS

DEFINITION Homo sapiens cDNA FLJ13331 fls, clone OVARC1001809, moderately similar to Mus musculus springosine kinase (SPHK1a) mRNA.

ACCESSION AK023393

VERSION AK023393.1 GI:10435311

KEYWORDS

SOURCE Homo sapiens ovary, tumor tissue cDNA to mRNA, clone_ljb:OVARC1 clone:OVARC1001809.

ORGANISM Homo sapiens

REFERENCE

AUTHORS

1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Watanuma,M., Hosofiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murkaml,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahatai,K., Masuko,Y., Minomiyu,K. and Iwayanagi,T.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2 (bases 1 to 1772)

Isogai,T. and Otsuki,T.

Direct Submission

Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan, cDNA full insert sequencing. Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science,

FEATURES

source

location/Qualifiers

1..1772

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="OVARC1001809"

/tissue_type="ovary, tumor tissue"

/clone_ljb="OVARC1"

/note="cloning vector: pME18SFL3"

350..1504

/note="unnamed protein product"

/codon_start=1

/protein_id="BAB14558.1"

/db_xref="GI:10435312"

/translation="MDPAGSGRGVLPRPCRVLLNPRGKGKALOLFRRSHVPLAE AEISFTLATERRNHARELVSEELGRMDALVNSGDLMEHYVNGMERPDETALQ KPLCSLPGASGNALADLNHAYGEVYVNEDELINCTLLCRLLSPPNNLSLHTASG LRLPSVLSLAWGFLADVDLESEKRYRLEEMRFTLTFLRLAALRTYRGLAYLPGRV GSKTPASPVVVOQGVDAHLVPLEEPYSHWTVPDEDFVLVALLSHSGSEMFAP MGRCAAGVHMLFYRAGYSRAMLRLFLAMEGRHMECEPYLYVYPVAPRLPDKG KGVFAVDGELMVSFAVQGVHPNPFMYVSCVEPRPPSMKPPQMPPEEPL"

BASE COUNT 285 a 566 c 585 g 336 t

ORIGIN

Alignment Scores:

Pred. No.: 1,28e-163 Length: 1772

Score: 1962.00 Matches: 375

Percent Similarity: 98.18% Conservative: 2

Best Local Similarity: 97.66% Mismatches: 7

Query Match: 97.37% Indels: 0

DB: 9 Gaps: 0

US-09-784-810a-2 (1-384) x AK023393 (1-1772)

Qy 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20

Db 350 ATGGATCCAGGAGGCGCGCCCGCGGCGTGTCCCGCGCGCTGCGCTGCTGCTG 409

Qy 21 LeuAsnProArgGlyGlyGlyValAlaLeuGlnLeuPheArgSerHisValGlnPro 40

Db 410 CTGAACCGCGCGCGGCGAGAGGCGCTTGCAGCTTCCGAGTCCGAGTGCAGCGCC 469

Qy 41 LeuAlaGlnAlaGlnLeuSerPheThrLeuMetLeuThrGluArgArgAsnHisAla 60

Db 470 CTTTGGCTGAGGCTGGAATCTCTTCCGCTGATGCTCAGTGCAGCGCGGACACCGC 529

Qy 61 ArgGluLeuValArgSerGluGluLeuGlnArgTyrPaspAlaLeuValMetSerGly 80

Db 530 CGGAGCTGTGCGGCTGCGAGAGCTGGCGGCTGCGAGCTGTGCTCATGCTCGGA 589

Qy 81 AspGluLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTyrGluThrAla 100

Db 590 GACGGCTGAGTGCAGAGAGTGTGTAACGGCTTCATGAGAGCGCGCTGAGTGGAGACCGC 649

Qy 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu 120

Db 650 ATCCAGAGCCCGCTGTGAGCTCCACAGAGCTTGGCAACGCGCTGCGACTTCTTG 709

Qy 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140

Db 710 AACCATATGCTGCTATGACAGAGTCAACATGAAGACCTCCACCACTGACGAGCTTA 769

Qy 141 LeuLeuCysArgProValLeuSerPheMetAsnLeuLeuSerLeuHisThrAlaSerGly 160

Db 770 TTGCTGTGCGCGCGCGCTGCTGACCCATGAACCTGCTGTGCAACAGCGCTTGGGG 829

Qy 161 LeuArgSerPheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeuVal 180

Db 830 CTGCGCTCTTCTCTGTCTCAGCTGCGCTGGCGCTTCACTTGCATGCTGAGTGCAGT 889

Qy 181 SerAspLysTyrArgArgLeuGlnGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200

Db 890 AGTGAGAGTATGCGCGTCTGCGGAGATGCGCTTCACTGCGAGCACTTCTCGCTG 949

[illegible]

gene	/cell_type="vascular endothelial"
CDS	1. .1173 /gene="SPHK" 19. .1173 /gene="SPHK" /codon_start=1 /product="sphingosine kinase" /protein_id="AA601980.1" /db_xref="GI:9909361" /translation="MDPAGSPRGVLPBPCRYVLNDRGSKALOLFRRHVOPLLA AEISFTMLTERRHARELYRBEVETGLMDALVYMSGGLHNEVNGIMEPRDETALQ KPLCSIDPAGSGLNLAASLINVAGYEVDGNTDLITCSLLCRILSPMLLSLHTASG LRLEFSVLISLAMGFADVDLSEKRYRRLGEMRFTLGTTLRLALRTYNGRLATLPCNV GSKTPASPVVVOGPDVAHLVPLEEPVSHMTYVDDDEYLVATLALYSSEFARV MGRCAAPVHMLFYVRAGVSRAMLRLRLAEKGMHNEEDRYLVYRVVAFRLRDKG KGMFADVGEMLVSEAVOYHNPYFMVSVCSVEPRPPMKPQDMRPEPL"
BASE COUNT	188 a 349 c 390 g 246 t
ORIGIN	
Alignment Scores:	
Prod. No.:	1.4e-163 Length: 1173
Score:	1959.00 Matches: 374
Percent Similarity:	98.18% Conservative: 3
Best Local Similarity:	97.40% Mismatches: 7
Query Match:	97.22% Indels: 0
DB:	Gaps: 0
US-09-784-810A-2 (1-384) x AF200328 (1-1173)	
OY	1 MetAsProAlaIaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
Db	19 ATGATATCAGAGGGGGGGCCCCGGGGGGTGTCCCGGGCCCTCGCTGTGGTCTGTG 78
OY	21 LeuAsnProArgGlyGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
Db	79 CTGAACCCGGCGGGCGGCAAGGCGCTTGACGCTCTTCGGAGTCAAGTCAAGCC 138
OY	41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluIleArgGlyAsnHisAla 60
Db	139 CTTTGGCTGAGGCTGAAATCTCTTCACGGCTGATGCTCACTGAGCGGGGACACAGCG 198
OY	61 ArgGluLeuValAlaArgSerGluGluLeuGlnGlyArgTrpAspAlaLeuValAlaMetSerGly 80
Db	199 CGGAGGCTGTGGCGGTGGAGAGACTGGCGCGGCGAGACGCTGTGGTGTATGTCTGGA 258
OY	81 AspGlyLeuMetHisGluValValAsnGlyLeuMetGluArgProAspTrpGluThrAla 100
Db	259 GACGGGCTGATGACACAGAGTGGTAAACGGGCTCAATGGAGCGGCTGACGGAGACCGCC 318
OY	101 IleGlnIysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaAlaSerLeu 120
Db	319 ATCAGAAAGCCCGCTGTGATAGGCTCCCAAGACGCTCTGGCAAGCGGCTGGCACTTCCTTG 378
OY	121 AsnHisIysAlaGlyIleGlyTrpGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
Db	379 AACCATATATCGTGGCTATGACACAGTCAACCATGAAACCTCTGACCAACTGCACCTTA 438
OY	141 LeuLeuCysArgProValLeuSerPheMetAsnLeuSerLeuHisThrAlaSerGly 160
Db	439 TTGCTGTCCCGCCGGGCTGTCTCAACCATGAACTGCTGTCTGCAACACGGCTTCGGGG 498
OY	161 LeuArgSerPheSerValLeuSerLeuAlaIleArgGlyPheIleAlaAspValAspLeuGlu 180
Db	499 CTGGCGCCTCTCTGTGTCTACAGCTGGCGCTGGGGCTTCATTTGCTGATGTGAGACTGAG 558
OY	181 SerAspIysIysIysArgGlyGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
Db	559 AGTAGAAGATATCGGCGCTGTGGGGAGATGCGCTTCACTGTGGGCACTTTCCTCGTCTG 618
OY	201 AlaAlaLeuAlaGlyTrpIysArgGlyArgGluAlaThrLeuProValGlyArgValGlyPhe 220
Db	619 GCACCCCTTGGCCACCTTACCGCGGCGGAGACTGGCTTACCTCCCTGTAGGAAGAGTGGGTTCC 678

QY 221 LysThrProAlaSerProValValAlaInGlnGlyProValAlaSpAlaHisLeuValPro 240
 DB 679 AAGACACCTGCTCCCTCCCGCTGGTGTGCAGAGGCGCCCGGTAGATGCACCTTTGTGCA 738
 QY 241 LeuGluGlnValProSerHisTrpGlnValAlaProAspGlnAspPheValLeuVal 260
 DB 739 CTGGAGGAGCAGTGCCTCTCAGTGCAGATGGTGCCGACGAGACTTTGTCTAGTC 798
 QY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGlnUrePheAlaAlaProMetGlyArgCys 280
 DB 799 CTGGACACTGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 858
 QY 281 AlaAlaGlyAlaMetHisLeuPheTyrValAlaArgAlaGlyAlaSerArgAlaMetLeu 300
 DB 859 GCAAGCTGGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 918
 QY 301 ArgLeuPheLeuAlaMetGlnUreGlyArgHisMetGlnUreGlyArgCysProTyrLeuVal 320
 DB 919 CGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 978
 QY 321 TyrValProValAlaAlaPheArgLeuGlnUreProLysAspGlyArgValAlaPheAlaVal 340
 DB 979 TATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
 QY 341 AspGlyGlnUreMetValSerGlnAlaValGlnGlnGlnValHisProAsnTyrPheTrp 360
 DB 1039 GATGGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1098
 QY 361 MetValSerGlyCysValGlnUreProProSerTrpLysProGlnGlnMetProProPro 380
 DB 1099 ATGTGACGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1158
 QY 381 GlnGluProLeu 384
 DB 1159 GAAGAGCCCTTA 1170
 RESULT 8
 LOCUS BC008040 1824 bp mRNA linear pri 12-JUL-2001
 DEFINITION Homo sapiens, clone MGC:15041 IMAGE:3831657, mRNA, complete cds.
 ACCESSION BC008040
 VERSION BC008040.1 GI:14165485
 KEYWORDS MGC.
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1824)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (21-MAY-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 CONTACT: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
 http://www.systemsbio.org
 contact: amadan@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettman and Anuradha Madan
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRL Plate: 24 Row: n Column: 4
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10433790.
 FEATURES
 SOURCE Location/Qualifiers
 1..1824

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:15041 IMAGE:3831657"
 /tissue_type="Skin, melanotic melanoma."
 /clone_lib="NIH_MGC_20"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 339..1493
 /product="Unknown (protein for MGC:15041)"
 /protein_id="AA08040.1"
 /db_xref="GI:14165485"
 /translation="MDPAGRGVLPAPRCRYVILNPRGKKAQILRSHVOPILAE
 AETSPTLMIFERNHARELVSSELDGRMDALVMSGGILHVEYMLERDPMETATQ
 KPLCSLPASSGNALASLNHYAGYEYVINDLLNCTLLICRLSLPMNLSTHPSAG
 LRLESVLSLNGEIVADVDESEKFRRLGEMKFTLGTEFLLAALRTYRGRALYLPVGRV
 GSKTPASPVVVOGEPVDAHLVLEPVSMTVVPDEPVLVALHSHLSGEMFAP
 MGRCAGVNHLFVRAVSRAMLRFLAMEKGRHMEYECPLYLVVYVAVFRLPKDG
 KGVFAVDGELMSEAVOCVHPNFYFMVSGCVPEPPSMKPOQMPPEEPL"
 BASE COUNT 354 a 551 c 592 g 327 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,49e-163 Length: 1824
 Score: 1955.00 Matches: 374
 Percent Similarity: 97.92% Conservative: 2
 Best Local Similarity: 97.40% Mismatches: 8
 Query Match: 97.02% Indels: 0
 DB: 9 Gaps: 0
 US-09-784-810A-2 (1-384) x BC008040 (1-1824)
 QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProArgValLeuValLeu 20
 DB 339 ATGATCATCAGCGGGGGCCCGGGGGCGTGTCCCGGGGCGCTGCGCGGTGGTGGT 398
 QY 21 LeuAspProArgGlyGlyGlyGlyValAlaGlnUrePheArgSerHisValGlnPro 40
 DB 399 CTGAACCG 458
 QY 41 LeuLeuAlaGlnUreGlySerPheTrpLeuMetLeuTrpGlnUreArgArgAsnHisAla 60
 DB 459 CTTTGGCTGAGGCTGAAATCTCTTCACTGATGCTGATGCTGATGCTGATGCTGATG 518
 QY 61 ArgGlnUreValArgSerGlnUreGlnUreGlnUreGlnUreGlnUreGlnUreGlnUre 80
 DB 519 CGGAGCTGTGTGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 578
 QY 81 AspGlyLeuMetHisGlnValValAlaSerGlyLeuMetGlnUreArgProAspTrpGlnUreAla 100
 DB 579 GACGGCTGTGATGCACAGAGGTGTGAAGCGGCTCATGAGGCGGCTGACGAGGAGCGCC 638
 QY 101 IlleGlnUreProLeuGlySerLeuProAlaGlySerGlyAlaAlaAlaAlaSerLeu 120
 DB 639 ATCCAGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
 QY 121 AsnHisTyrAlaGlyTyrGlnGlnValAlaTrpAsnGlnUrePheTrpAsnGlyTrpLeu 140
 DB 699 AACCATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
 QY 141 LeuLeuGlyArgProValLeuSerProMetLeuLeuLeuSerLeuHisTrpAlaSerGly 160
 DB 759 TTGCTGTGGCGCGCGCTGTGTACCATGCACTGCTGTCTGTGCAACGGGCTTGTGGGG 818
 QY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAlaSerLeu 180
 DB 819 CTGGCGCTCTTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 878
 QY 181 SerAspLysTyrArgArgLeuGlyGlnUreArgPheTrpLeuGlyTrpPheLeuArgLeu 200
 DB 879 AGTAGAAGATATGGGCTGTGGGAGAGATGGGCTTCACTGCGGACACTTCTCGCTTCG 938
 QY 201 AlaAlaLeuArgTrpTyrArgGlyArgLeuAlaThrLeuProValAlaGlyArgValGlyPhe 220

Db 981 CTGGGGAGATGCCCTTCACTCTGGGACCTTCTGGCTGGGACCGCTGCGACCTAC 1040
 Oy ArgglyArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrProAlaSerPro 226
 Db 1041 CGCGCGGACGAGCTGCTCTCCCTGTGAGAGAGTGGGTTCACAGACCTGCTCCGCC 1100
 Oy 227 ValValValGlnGlnGlnGlnProValAlaPalaHisLeuValProLeuGlnGlnValPro 246
 Db 1101 GTTGTGTCCAGCAGGCGCGGTGATGACACACTTGGCCACCTGAGAGACCACTGCGCC 1160
 Oy 247 SerHisTrpGlnValValProAspGluAspPheValLeuValLeuAlaLeuHisSer 266
 Db 1161 TCTCCTGACAGAGTGGTCCCGACGAGAGACTTGTGCTAGTCTGCGACCTGCTCCTCG 1220
 Oy 267 HistLeuAlaSerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGlnValMetHis 286
 Db 1221 CACCTGGGACGCTGAGAGTGTGCTGACACCCAGGCGCGCTGTGACGCTGGGCTCATGCT 1280
 Oy 287 LeuPheTrpValAlaGlnValAlaSerArgAlaMetLeuLeuAlaGluPheLeuAlaMet 306
 Db 1281 CTGTTCTACGTGGGGGGGAGTGTCTGTCGTCATGCTGCGCTGCGCTCTCTGCGCATG 1340
 Oy 307 GluLysGlyArgHisMetGlyTrpGlyCysProTrpLeuValTrpValProValAla 326
 Db 1341 GAGAGGGGACGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1400
 Oy 327 PheArgLeuGluProLysAspGlyLysGlyValPheAlaValAspGlyLysLeuMetVal 346
 Db 1401 TTCCGCTTGGAGCCCAAGATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1460
 Oy 347 SerGluAlaValGlnGlnGlnValAlaHisProAsnTrpPheTrpMetValSerGlyCysVal 366
 Db 1461 AGCAGAGGCGCGGAGGCGGAGGTGACACCAACTACTCTGTGATGTGTGTGTGTGTGTG 1520
 Oy 367 GluProProProSerTrpLysProGlnGlnMetProProProGluGluProLeu 384
 Db 1521 GAGCCCGGCGGACCTGAGAGCCCGGACGATGCGCACCGCAGAGAGCCCTTA 1574
 RESULT 11
 AK095578 2502 bp mRNA linear PRI 15-JUL-2002
 LOCUS Homo sapiens cDNA FLJ38259 f1s, clone FCBBF3001302, highly similar
 DEFINITION to Homo sapiens sphingosine kinase (SPHK1) mRNA.
 ACCESSION AK095578
 VERSION AK095578.1 GI:21754862
 KEYWORDS oligo capping: f1s (full insert sequence).
 SOURCE Homo sapiens fetal brain cDNA to mRNA, clone_lib:FCBBF3
 Clone:FCBBF3001302.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,
 Isono, Y., Kawai, H., Y., Saito, K., Miki, K., Miki, K., Kikuchi, H.,
 Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kimura, K.,
 Kanda, K., Magatsuma, M., Murakawa, K., Kanehori, K.,
 Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K.,
 Masuho, Y., Nagai, K. and Isogai, T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2502)
 AUTHORS Isogai, T. and Yamamoto, J.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUL-2002) Takao Isogai: FLJ Project(HRI Team): 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan: cDNA full insert sequencing:
 Research Association for Biotechnology (RAB): cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.): 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 source
 1..2502
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="FCBBF3001302"
 /tissue_type="brain"
 /clone_lib="FCBBF3"
 /dev_stage="fetal"
 /note="Cloning vector: pME18SFL3"
 BASE COUNT 404 a 803 c 835 g 460 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.65e-162 Length: 2502
 Score: 1944.00 Matches: 372
 Percent Similarity: 98.16% Conservative: 2
 Best Local Similarity: 97.64% Mismatches: 7
 Query Match: 96.48% Indels: 0
 DB: 9 Gaps: 0.
 US-09-784-810a-2 (1-384) x AK095578 (1-2502)
 Oy 4 AlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeuLeuAsnPro 23
 Db 1087 GCGGGGCGGCGCGGGGGGCTCTCCCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1146
 Oy 24 ArgGlyGlyGlyGlyLysAlaLeuGlnLeuPheArgSerHisValGlnProLeuLeuAla 43
 Db 1147 CGCGCGGCAAGGCGCAAGGCGCTTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 1206
 Oy 44 GluAlaGluLeuSerPheTrpPheTrpLeuMetLeuTrpGluArgArgAsnHisAlaArgGluLeu 63
 Db 1207 GAGGCTGAATCTCTTACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266
 Oy 64 ValArgSerGluGluLeuGlyArgTrpAspAlaLeuValAlaMetSerGlyAspGlyLeu 83
 Db 1267 GTGCGGTGCGAGAGAGCTGGCGGTGGAGAGCTGCGGTGGAGAGCTGCGGTGGAGAGAGCTG 1326
 Oy 84 MethisGluValValAsnGlyLeuMetCysLysArgProAspTrpGluThrAlaIleGlnLys 103
 Db 1327 ATGACAGAGTGTGAGAGGCGCTCATGAGAGGCGCTGAGAGAGCGCATCCAGAGAG 1386
 Oy 104 ProLeuCysSerLeuProAlaGlySerGlyAsnAlaAlaSerLeuAsnHisTrp 123
 Db 1387 CCCCTGTACCTCCAGCAGAGCTGTGCAACGCGCTGCAACGCTGTGCAACGCTGTGCAAC 1446
 Oy 124 AlaGlyTrpGlnGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeuLeuLeuCys 143
 Db 1447 GCTGGCTATGAGCAGAGGTGACCAATGAAAGACCTCGTACCAAGTGTGCTGCTGCTGCTG 1506
 Oy 144 ArgProValLeuSerProMetAsnLeuLeuSerLeuHisTrpHisAlaSerGlyLeuArgSer 163
 Db 1507 CGCGGCGTGTGCTACCATTAACCTGTGCTGTGCAACGCGCTGTGCAACGCGCTGTGCAAC 1566
 Oy 164 PheSerValLeuSerLeuAlaTrpGlyPheLeuAlaAspValAspLeuGluSerAspLys 183
 Db 1567 TTCTCTGTGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1626
 Oy 184 TyrArgArgLeuGlyGluMetArgPheTrpLeuGlyThrPheLeuArgLeuAlaLeu 203
 Db 1627 TATGCGGCTGTGGGGGAGATGCGCTTCACTGCGGACCTTCTGCGTGGAGCGCTG 1686
 Oy 204 ArgThrTrpArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPheLysThrPro 223
 Db 1687 CGCAGCTACCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1746
 Oy 224 AlaSerProValValAlaGlnGlnGlnProValAlaPalaHisLeuValProLeuGln 243
 Db 1747 GCTCTCCCGGTTGT 1806
 Oy 244 GlnValProSerHisTrpGlnValValProAspGluAspPheValLeuValLeuAlaLeu 263

DB 1807 CCAGTCCCTCTCAGTGGACATGTGGTGGCCGAGGAGACTTGTGCTAGTCCCTGACCTG 1866
QY 264 LeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCysAlaAlaGly 283
DB 1867 CTCGACTCGCAGCTGGGAGATGATGTTGCTGACCCATGGCCGCTGTGACAGTGC 1926
QY 284 ValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeuArgLeuPhe 303
DB 1927 GTCTAGCATCTGTCTTACGTCGGGGGGAGTGTCTGTCGATCTGCTGCGCTCTTC 1986
QY 304 LeuAlaMetGluLysGlyArgHisMetGluTyrGluCysProTyrLeuValTyrValPro 323
DB 1987 CTGGCATGAGAGAGGAGGAGCATATGAGATATGATGATGCTACTTGGTATACGTGGCC 2046
QY 324 ValValAlaPheArgLeuGluProLysAspGlyLysGlyValPheAlaValAspGlyGlu 343
DB 2047 GTGGTCCCTCTCCCTTGGAGCCCAAGATGGGAAAGTGTCTTCTCAGTGGATGGGAA 2106
QY 344 LeuMetValSerGluAlaValAlaGlnGlyGlnValHisProAsnTyrPheTyrMetValSer 363
DB 2107 TTGATGGTTACCGAGGCGGTGACAGGCGAGTGCACCCAACTACTTGTGATGTCAGC 2166
QY 364 GlyCysValGluProProSerTyrPlySerProGlnMetProProProGluGluPro 383
DB 2167 GGTTCGTGGAGCCCGCCAGCTGGAGAGCCCGAGAGATGCGACCGCCAGAGAGGCC 2226
QY 384 Leu 384
DB 2227 TTA 2229

RESULT 12
HSA245504 1428 bp mRNA linear PRI 22-MAY-2000
LOCUS HSA245504
DEFINITION Homo sapiens mRNA for sphingosine kinase (gene SPK).
ACCESSION AJ245504
VERSION AJ245504.1 GI:8017375
KEYWORDS sphingosine kinase; SPK gene.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1428)
AUTHORS Van Veldhoven, P.P. and Gijssbers, S.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1428)
AUTHORS Van Veldhoven, P.P.
TITLE Direct Submission
SUBMITTED (16-AUG-1999) Van Veldhoven P.P., Molecular Cell Biology,
Katholieke Universiteit Leuven, K.U. Leuven, Campus Gasthuisberg,
Afd. Farmakologie, Herestraat, B-3000 Leuven, BELGIUM
REMARK REVISD BY author (17-MAY-2000)
FEATURES
Source
1..1428
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q24"
1..1146
/gene="SPK"
1..1146
/gene="SPK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="CAB92131.1"
/translation="MDPAGPRGVARLRLVILNLRGKGRKALQFRSHVQPLAERI
STLMTERRNNAELVSEELGRMDALVYVNGIMHEVNDIETAIQKPL
CSLPAGSGPLASLNTYAGTEQVLTNDLLNCLRLSPMNLISLSTASGLT
FVSLAMGFIADVLESEKTRLEGEKFTGCTELRLAATRTYGRALAYIPVAGSGK
TPASPVVQGGFADVAHVLPLEPVSHTVVPDEDTIYVLAALSHSGSMFAPAKR
CAAGVHLTYVRAGVRAMLRLFLAMKGRHMEYECTLVITVYGAFLRLEPKRKY
PAVDGLMAYSEAVGCVHNTFMVNSGCVPEPPSWKPOAMPPEPL"

BASE COUNT 247 a 422 c 462 g 297 t
ORIGIN

Alignment Scores:
Pred. No.: 1, 5e-156 Length: 1428
Score: 1880.50 Matches: 366
Percent Similarity: 95.838 Conservative: 2
Best Local Similarity: 95.318 Mismatches: 13
Query Match: 93.334 Indels: 3
DB: 9 Gaps: 2

US-09-784-810a-2 (1-384) x HSA245504 (1-1428)

QY 1 MetAspProAlaGlyGlyProArgGlyValLeuProArgProCysArgValLeuValLeu 20
DB 1 ATGATCTCAGCG--GGGCCCCGGGGCGTT-----GCTCGCCCTGCGGCTGTGCTG 51
QY 21 LeuAsnProArgGlyGlyLysGlyLysValAlaLeuGlnLeuPheArgSerHisValGlnPro 40
DB 52 CTGAACCCGCGCGGCGGCAAGGCAAGGCGCTTGCACCTCTCGAGAGTCACGTGACGCC 111
QY 41 LeuLeuAlaGluAlaGluIleSerPheThrLeuMetLeuThrGluArgAsnHisAla 60
DB 112 CTTTGGCTGAGCTGAAATCTCTTACGCTGATGCTCAGTACGCGGGAACACGCG 171
QY 61 ArgGluLeuValArgSerGluGluLeuGlyArgTyrPheAspAlaLeuValMetSerGly 80
DB 172 CGGAGCTGTGCGGTGAGAGAGCTGAGAGCTGGCGGCTGGAGACCTTGTTGTTGTA 231
QY 81 AspGlyLeuMetHisGluValValAlaAsnGlyLeuMetGluArgProAspTyrGluThrAla 100
DB 232 GAGGGCTGATGACAGAGAGTGTGTAAGCGGCTCATGGAGGCGCTGACTGGAGACCGCC 291
QY 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaAlaSerLeu 120
DB 292 ATCCAGAAAGCCCTGTGTAGCTCCACAGAGCTGTGGCAACCGCTGGCAGCTTCTTG 351
QY 121 AsnHisTyrAlaGlyTyrGluGlnValThrAsnGluAspLeuLeuThrAsnCysThrLeu 140
DB 352 AACCATTAAGCTGCTGCTATGACAGGTACCAATGAAAGCATCTGACCACTGACGCTA 411
QY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuLeuSerLeuHisThrAlaSerGly 160
DB 412 TTGCTGCGCGCGGCTGCTGTCACCAAGAACTGCTGCTCTCTCCACAGCGTTCGGGG 471
QY 161 LeuArgSerPheSerValLeuSerLeuAlaTyrGlyPheIleAlaAspValAspLeuGlu 180
DB 472 CTGGCGCTTCTCTCTGCTGCTGACCGCTGGCGCTGCTGATGTGACTAGAG 531
QY 181 SerAspLysTyrArgArgLeuGlyGluMetArgPheThrLeuGlyThrPheLeuArgLeu 200
DB 532 AGTGAGAAGATGCGGCTGTGGGGAGATGCGCTTACTCTGCGACCTTCCTGCGCTG 591
QY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValAlaPhe 220
DB 592 GCAGCCCTCGCCACCTACCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
QY 221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro 240
DB 652 AAGACACCTGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
QY 241 LeuGluGluGlnValProSerHisTyrPheValValProAspGluAspPheValLeuVal 260
DB 712 CTGGAGAGGCCAGTGCCCTCTCACTGAGAGTGAGTCCGACAGAGAGACTTGTGCTAGTC 771
QY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGluMetPheAlaAlaProMetGlyArgCys 280
DB 772 CTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 831
QY 281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu 300
DB 832 GCAGCTGCGTATGATCTGTCTTACTGTGGGGGGAGAGTGTCTGTCGATGCTGCTG 891

	*	53623	56955: contig of 3333 bp	in length
	*	56956	57055: gap of 100 bp	
	*	57056	59555: contig of 2500 bp	in length
	*	59556	59655: gap of 100 bp	
	*	59656	62713: contig of 3038 bp	in length
	*	62714	62813: gap of 100 bp	
	*	62814	66519: contig of 3706 bp	in length
	*	66520	66619: gap of 100 bp	
	*	66620	66930: contig of 2311 bp	in length
	*	66931	69030: gap of 100 bp	
	*	69031	72724: contig of 3694 bp	in length
	*	72725	72824: gap of 100 bp	
	*	72825	76092: contig of 3266 bp	in length
	*	76093	76192: gap of 100 bp	
	*	76193	80143: contig of 3951 bp	in length
	*	80144	80243: gap of 100 bp	
	*	80244	84957: contig of 4714 bp	in length
	*	84958	85057: gap of 100 bp	
	*	85058	86424: contig of 3367 bp	in length
	*	86425	86524: gap of 100 bp	
	*	86525	93428: contig of 4904 bp	in length
	*	93429	93528: gap of 100 bp	
	*	93529	97177: contig of 3649 bp	in length
	*	97178	97277: gap of 100 bp	
	*	97278	102713: contig of 5436 bp	in length
	*	102714	102813: gap of 100 bp	
	*	102814	109141: contig of 6328 bp	in length
	*	109142	109241: gap of 100 bp	
	*	109242	114513: contig of 5272 bp	in length
	*	114514	114613: gap of 100 bp	
	*	114614	124525: contig of 9912 bp	in length
	*	124526	124625: gap of 100 bp	
	*	124626	132364: contig of 7733 bp	in length
	*	132365	133464: gap of 100 bp	
	*	133465	139488: contig of 7024 bp	in length
	*	139489	139588: gap of 100 bp	
	*	139589	148734: contig of 9146 bp	in length
	*	148735	148834: gap of 100 bp	
	*	148835	159465: contig of 10631 bp	in length
	*	159466	159565: gap of 100 bp	
	*	159566	172885: contig of 13120 bp	in length
	*	172886	172885: gap of 100 bp	
	*	172886	183443: contig of 10658 bp	in length
FEATURES		Location/Qualifiers		
	source	1..183443		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/chromosome="17"		
		/map="17"		
		/clone="CTD-2531H7"		
		/clone_1fb="CYTD Hunan BAC"		
	misc_feature	1..1059		
		/note="assembly-fragment"		
misc_feature		1160..2777		
		/note="assembly-fragment"		
misc_feature		2878..4283		
		/note="assembly-fragment"		
misc_feature		4384..5447		
		/note="assembly-fragment"		
misc_feature		5348..6776		
		/note="assembly-fragment"		
misc_feature		6877..8215		
		/note="assembly-fragment"		
misc_feature		8316..9961		
		/note="assembly-fragment"		
misc_feature		10063..12224		
		/note="assembly-fragment"		
misc_feature		12325..13794		
		/note="assembly-fragment"		
misc_feature		13895..16857		
		/note="assembly-fragment"		
misc_feature		16936..18266		
		/note="assembly-fragment"		

	misc_feature	18367..*20190 note="assembly_fragment"
	misc_feature	20291..21460
Alignment Scores:		
Pred. No.:	6.28e-148	Length: 183443
Score:	1813.00	Matches: 371
Percent Similarity:	69.85%	Conservative: 2
Best Local Similarity:	69.48%	Mismatches: 8
Query Match:	89.98%	Indels: 154
DB:	2	Gaps: 3
US-09-784-810A-2 (1-384) x AC068145 (1-183443)		
OY	4	AAlAglygylProArGyglValleuProApqPocysArqValleuValleuDeuansPro 23
Db	135039	GGGGGGGGCCCCCGGGCGTGTCTCCCGGGGCCCTGGCCGCTGTGTGTGTGCAGACCG 134980
OY	24	ArGygylgylVsgylVtlysAlaleuGlnleuPharArgSerHstValjInProleuenua 43
Db	134979	CGGCGCGCAGAGGGCAAGSGCTTGACGCTTCGCGAGTCACTGCAGCCCTTTTGCT 134920
OY	44	GlulAlagluIleSerphrLeuMetleuPr----- 54
Db	134919	GAGGCTGAATCTCTTCACGCTGATGCTACTGGTAGTACTCTCGAGGGGTTTTGG 134860
OY	55	-----Glu 55
Db	134859	GGAGCATCCCTGGCACGGGAGCCCCCCAGCTCTGATAGCTGCCGNTCTCTTCAGAG 134800
OY	56	ArgArGasnHlSaIAARgIleuValArgSerIugIleuGlyArgrTPAsPalaleu 75
Db	134799	CGCGGAAACCAACGGCGSAGCTGTGGCTGGAGAGAGCTGGGCCGCTGGAGCCTCTG 134740
OY	76	VAlAlmetSerGIyASpgIyleuMeThSLyuAl----- 87
Db	134739	GTGGTCAATGTCTGGAGAGAGGGGTGATGACACAGAGT-GAGAACCGCACTGCCGCTAGC 134681
OY	88	-----ValAsn 89
Db	134680	CTGGGGCTTGGCGCGGTGCTCCAGAGCTGAGAGCCAGCTGTCTTCACAGSTGGTGAAC 134621
OY	90	GlyLeuMeTgluarProAsprTPsgIuthralalIegIuNysProLecycSerleuPro 109
Db	134620	GGGCTCATGGAGCGGCGCTGTGACTGGAGACCGGCATCCAGAAGGCCCTGTGTAGCTCCA 134561
OY	110	AlAgIyserGIyASnalaleuAlaleuAlaseleuAnsHlstryAlagly----- 125
Db	134560	GCAGGCTCGGCAACCGCTGGCAGCTTCTTGAACATTATGTCTGGTGAAGGCCAGG 134501
OY	125	----- 125
Db	134500	GCCAGAGTAVAGGCGTGTTCGCCGTCAATGCTCTCTACCGCGGGGTTTTCTTGCTAAGC 134441
OY	125	----- 125
Db	134440	TCCCATAGSGGTGAGATCATTTTCATTTCCCCTTCTCCCTTAGTCGCTGGGGCCCTCTCG 134381
OY	125	----- 125
Db	134380	GTGACCCAGCTGACTGCTTTCATTTGTCTTCATCTGTACCTTACCTACAGTCTGCCAATCC 134321
OY	125	----- 125
Db	134320	CCAGGAGCACATGAGCGCTTGGCAGACTCCACTCCCGGAGAGAGAACCGGGGATAC 134261
OY	126	-----TyrluJInValThr 130
Db	134260	ATGGGGGCTCCTGTCTCTTATCTGCACTTTTCCCGCTGACAGCTATGAGCAGGTACC 134201
OY	131	AsngIuAspLeuIethrAsnyCystrIleuIeuIenCysArgProValleuSerPromet 150
Db	134200	AATGAAGACTCTGTGACCACTGACAGCTATTGGCTGTGCGCGCGCTGTACCCATYG 134141

Oy	151	IsnleuLeuSerLeuHisThrAlaSerGlyLeuArgSerPheSerValLeuSerIleuAla	170
Db	134140	AACCGCTGTCTCTGCACACGGCTTCGGGGCTGGCGCTCTTCTCTGTGTCAACCGGGCC	134081
Oy	171	TrpGlyPheIleAlaAspValAspLeuGluSerAspLysTyrArgArgLeuGlyGluMet	190
Db	134080	TGGGGCTTCATTGCTGCATGTGTGACCTAGAGACTAGAAATATACGGCTCTGGGGAGATG	134021
Oy	191	ArgPheThrLeuGlyThrPheLeuArgLeuAlaIleuArgThrTyrArgGlyArgLeu	210
Db	134020	CGCTTCACTCTGGGGCACCTTCTCGCTGTGGCAGCCCTCGCACACTACCGCGCGCAGTG	133961
Oy	211	AlaThrIleuProValGlyArgValGlyPheLysThrProAlaSerProValValAlaGln	230
Db	133960	GCCATACCTCCCTGTAGGAAGATGGTGTCCAAAGACACCTGGCTCCCGCTGTGGTCCAG	133901
Oy	231	GlnGlyProValAspAlaHisIleuValProLeuGlnGluGlnValProSerHisThrPglu	250
Db	133900	CAGGGCCCGGTAGATGACACACTTGTGCCACTGTGGAGAGACCACTGGCCCTCTCACTGGACA	133841
Oy	251	ValValProAspGluAspPheValLeuValLeuAlaLeuIleuHisSerHisLeuAlaSer	270
Db	133840	GTGGGCCCCGACAGAGACTTTGTGTAGCTGTGGCACTGTGCACCTGCACACCTGGCGAGT	133781
Oy	271	GluMetPheAlaAlaProMetGlyArgCysAlaAlaGlyValMethIleuPheTyrVal	290
Db	133780	GAGATGTTTGCTGCACCCACATGGGGCGCTGTGGACGTGGCGTCAATGATCTGTTCTACGTG	133721
Oy	291	ArgAlaGlyValSerArgAlaMetLeuLeuArgGlyPheLeuAlaMetGlyArgGlyArg	310
Db	133720	CGGGCGGAGTGTCTGTGGCCATGCTGCTGGCCCTTCTCTGGCCATGGAGAAAGGGCAGG	133661
Oy	311	HisMetGlyTyrGlyCysProTyrIleuValTyrValProValValAlaPheArgLeuGlu	330
Db	133660	CATAGGAGTATGAATGCCCTCACTGTGTATGTGCCGTGTGTCTTCCCTTCCCTTGAG	133601
Oy	331	ProLysAspArgGlyGlyValPheAlaValAlaAspGlyGlyIleuMetValSerGluAlaVal	350
Db	133600	CCCAAGATGGGAAAGTGTGTTCACAGTGGATGGGGAATTGATGGTTAGCGAGGCCGTG	133541
Oy	351	GlnGlyGlnValHisProAsnTyrPheTrpMetValSerGlyCysValGluProPro	370
Db	133540	CAGGGCCAGGTGCACCCAAACTACTTGTGATGTGACGTGTGCTGTGGAGCCCCGGCC	133481
Oy	371	SerTrpLysProGlnGluMetProProGlnGluProLeu	384
Db	133480	AGCTGGAAGCCCCACAGATGCCACCGCCAGAAAGACCCCTTA	133439
RESULT 14			
AB046025			
LOCUS	AB046025	2015 bp	mRNA linear PRI 18-JUL-2000
DEFINITION	Macaca fascicularis brain cDNA, clone:QCEE-16602.		
ACCESSION	AB046025		
VERSION	AB046025.1	GI:9280105	
KEYWORDS	fts (full insert sequence).		
SOURCE	Macaca fascicularis adult cDNA to mRNA, clone_lib:macaque brain cDNA library Qcee clone:QCEE-16602.		
ORGANISM	Macaca fascicularis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.		
AUTHORS	1 (sites)		
TITLE	Tseda, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M., Terano, K., Suzuki, Y., Sugano, S. and Hashimoto, K.		
JOURNAL	Isolation of full-length cDNA clones from macaque brain cDNA libraries		
REFERENCE	2 (bases 1 to 2015)		
AUTHORS	Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-JUL-2000) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama		

COMMENT

1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: hashimoto.h.go.jp; URL: http://www.nih.go.jp/yoken/genebank/,
Tel.: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
URL: http://www.nih.go.jp/yoken/genebank/
Lib Name: macaque brain cdna library Qcce
Lab host: TOP10
Vector: pME18S-FL3 (Acc. No. AB009864)
R. Site1: DraIII (CACCTGTC)C
R. Site2: DraIII (CACCATGTC)T
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCCTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by Sugano et al. (University of Tokyo, Institute of
Medical Science). Custom primer used for sequencing
(5' end primer [CTTCTGCTTAAGACTGC];
3' end primer [CGACCTGCAGCTCGACCA]);
Location/Qualifiers

FEATURES

source

1..2015
/organism="Macaque fascicularis"
/db_xref="taxon:9541"
/clone="QcCE-16602"
/clone_lib="macaque brain cdna library QcCE"
/dev_stage="adult"
628..1617
/codon_start=1
/product="unnamed protein product"
/protein_id="BAB01607.1"
/db_xref="GI:9280106."
/translation="MDPAEGRGVLPSCRVLVLLNPGKGKALQDFRSYHVDLAEE
AEISFTMLTERRHARELVSEELGRDADLVNMGGLMEVNGLMERDMETAIQ
KLCLSPAGSGNALAASLNHYAGEOYTNEEDILNCRLCLRSLPNMLLSHTGA
LRFSVSLANGFLADVDESEKRYRGRSEMFGTLGTRALSTYGSRALAYIPVGRA
GSKTPSVSYVYQGPRVDNHLPILEPPPSHHVTVPDDFVYLALLSHLSGEFAPAR
MGRCAAGVTHLFYRVAGYSRAMLRFLTEMERGNHMETECPTLYITIGWGDG-

CDS

BASE COUNT 355 a 639 c 650 g 371 t
ORIGIN

Alignment Scores:

Score: 2.48e-148 Length: 2015
Pident: 1789.50 Matches: 349
Percent Similarity: 91.43% Conservative: 15
Best Local Similarity: 90.65% Mismatch: 3
Query Match: 88.81% Indels: 18
DB: 9 Gaps: 1

US-09-784-B10A-2 (1-384) x AB046025 (1-2015)

Qy 1 MetaspproalaglgllyProargglyalleuproarprocyarsgvalleuValleu 20
Db 628 ATGATCTCCAGGGGGGGCCCCCGGGGGCTCTCCCGGGTCTGCGGGTGCTGCTG 687
Qy 21 Leuasnproargglylylysallylsalauleuglnleupeargserhisvalgnfro 40
Db 688 CTGAACCCGGCGGGGCCAGAAGGAGGCCCTGAGCTCTCCGGAGTACAGTGCAGCCC 747
Qy 41 leutaulagluallagluilleserphenrileumetleuthrgluatgatAsnHisAla 60
Db 748 CTTTGGGTGAGCGGAAATCTCTTCACGCTGATGAGCTCACTGAGCGGGGAGAACATTGCG 807
Qy 61 ArggluleuValArgserglugluleucllylrqtrppapalaleuValMetSerly 80
Db 808 CGGGAGCTGGTGGCTCGGAGAGCTGGGGCCGCTGGAGACCTCTGTGTGTATGTCTGGA 867
Qy 81 AspglyleuMetHisgluValValasnlyleumetgluarproaspfrpGlutHraLa 100
Db 868 GACGAGCTGATGACAGAGAGTGTGTAACGGGCTCATAGGAGCGGCTGACTGGAGACCGCC 927
Qy 101 lleglnlpSproLeucysSerieuProalaglyserglyasnallaleuAlaAlaSerieu 120

Db 928 ATCCAGAAAGCCCTGTGTAGCTCCAGACAGCTCTGGCAACCGCTGCACCTTCCTTG 987
 QY 121 AsnHisTyrAlaGlyTyrGlnGlnValThrAsnGluAspLeuThrAsnCysThrLeu 140
 Db 988 AACCATTAATGCTGCTATGAGCAGGTCACCAATGAACATCTCTGACCACTGCACGGGG 1047
 QY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160
 Db 1048 TTGCTGTCCCGCCGCTGCTACCTGACCAATGAACTGCTGCTGCTGACACGGCTTCGGGG 1107
 QY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu 180
 Db 1108 CTGGCCCTCTCTCTGCTGCTAAGCCCTGGCCCTGGGGCTTCAATGCTGATGGAGCTTAAG 1167
 QY 181 SerAspLysTyrArgArgLeuGlnGlyLysMetArgPheThrLeuGlnGlyThrPheLeuArgLeu 200
 Db 1168 AGTGAGAAAGTATCGCCCTGCGCCCTGCGGAGATGCCCTTCACTGCGGACCTTCCTGCTG 1227
 QY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
 Db 1228 GAGGCCCTGCACCTTACCTGAGCCGACCTGCTACCTCCTGTAGAGAGAGCGGCTCT 1287
 QY 221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro 240
 Db 1288 AAGACACCTGTTCCCTCCCTGTGGTGCAGAGGCCCCCGTAGATGCACACCTGTGGCA 1347
 QY 241 LeuGlnGlnGlnValProSerHisTyrPglValValProAspGluAspPheValLeuVal 260
 Db 1348 TTGGAGAGAGCAGGCTCTCTCACTGACAGAGTGGCCCGCACAGATTTTCGTCGGTG 1407
 QY 261 LeuAlaLeuLeuHisSerHisLeuAlaSerGlnMetPheAlaAlaProMetGlyArgCys 280
 Db 1408 CTAGGCGCTGCACCTGCACCTGCGGCGAGCGAGANTTTTGGCGACCACTGGCCGCTGT 1467
 QY 281 AlaAlaGlyValMetHisLeuPheTyrValArgAlaGlyValSerArgAlaMetLeuLeu 300
 Db 1468 GAGCTGTGTCAAGCTCTGCTCTATGTGTGCGGGGAGAGTGTCTGGCGCATGGCTGCG 1527
 QY 301 ArgLeuPheLeuAlaMetGlnGlyGlnGlyArgHisMetGlnTyrGlnCysProTyrLeuVal 320
 Db 1528 CCCCCTTCTCTGCGCAGAGAGGAGGACACATAGATGAATGATGCCCTTCACTGTGTA 1587
 QY 321 TyrValProValValAlaPheArgLeuGlnProLysAspGlyLysGlyValPheAla-Val 340
 Db 1588 TAC-----ATCAGT 1596
 QY 340 LaspGlyLysLeuMetValSerGlnAlaValGlnGlyValHisProAsnTyrPheTr 360
 Db 1597 GATGGGGAATTTGATGTTAGCGAGGCTGTGACAGGCCAGGTGACCAACATCACTTGG 1656
 QY 360 PheTyrAlaSerGlyCysValGlnProProSerTyrLysArgGlnGlnMetProProPr 380
 Db 1657 GATGGTCAATGATTTGGCTGGAGGCCCTGCCAGCTGAGAGCCCAAGCAGCGCACCTTC 1716
 QY 380 GclGlnProLeu 384
 Db 1717 AGAAGAGCCCTTA 1729
 RESULT 15
 AX224381 1759 bp DNA linear PAT 10-SEP-2001
 LOCUS AX224381 Sequence 3 from Patent WO0160990.
 DEFINITION AX224381
 ACCESSION AX224381
 VERSION AX224381.1 GI:1554632
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 1759)
 AUTHORS Rastelli, U.
 TITLE Novel sphingosine kinases
 JOURNAL Patent: WO 0160990-A 3 23-AUG-2001;

FEATURES Curagen Corporation (US) ; GENENTECH, INC. (US)
 Location/Qualifiers
 source 1..1759
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 BASE COUNT 357 a 495 c 523 g 384 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,32e-131 Length: 1759
 Score: 1594.50 Matches: 303
 Percent Simlarity: 88.25% Conservative: 35
 Best Local Simlarity: 79.11% Mismatches: 44
 Query Match: 79.13% Indels: 1
 Ds: 6 Gaps: 1
 US-09-784-810a-2 (1-384) x AX224381 (1-1759)
 QY 1 MetAspProAlaGlyGlyProAlaGlyValLeuProAlaGlyProCysArgValLeuValLeu 20
 Db 327 ATGGAACAGTAAATGCTCGAGAGCTCTCCACGGCCATGCAAGATGCTGTGCTG 386
 QY 21 LeuAsnProAlaGlyGlyLysGlyLysAlaLeuGlnLeuPheArgSerHisValGlnPro 40
 Db 387 CTGAACCCCGGAGGCGAGGCGAGGCGAGGCTGTGCACTCTTCACAGCCGCTGTGAGGCC 446
 QY 41 LeuLeuAlaGlnAlaGlnLysSerPheThrLeuMetLeuThrGlnGlnArgArgAsnHisAla 60
 Db 447 TTCTCGAGAGAGGACAGATGATACCTTAACCTGATACCTACCAAGAGAACCATGCC 506
 QY 61 ArgGlnLeuValArgSerGlnGlnGlnGlyArgTyrPAspAlaLeuValValMetSerGly 80
 Db 507 AGGAGCTGTGTGTGCGAGAGAGATTGGTCACTGGAGACCCCTGGCAGTCATGTCCGT 566
 QY 81 AspGlyLeuMetHisGlnValValAsnGlyLysMetGlnLysArgProAspTrpGlnTrpAla 100
 Db 567 GATGCTGTATCATGAAGTGTCAATGAGCTTATGAGACCCAGACCTGGAGACTGCC 626
 QY 101 IleGlnLysProLeuCysSerLeuProAlaGlySerGlyAsnAlaLeuAlaAspLeu 120
 Db 627 ATCCAAACACCCCTGTGAGCTCTCCGAGGCTCCGGAGTCCGAGATGGCGGACCTTCTG 686
 QY 121 AsnHisTyrAlaGlyTyrGlnGlnValThrAsnGluAspLeuThrAsnCysThrLeu 140
 Db 687 AACCATTAATGCTGCTATGAGCAGGTCACCAATGAACATCTCTGACCACTGCACACTG 746
 QY 141 LeuLeuCysArgProValLeuSerProMetAsnLeuSerLeuHisThrAlaSerGly 160
 Db 747 CTGTGTGCGCGCGCGCTGTACCACTGACCACTGCTGTGCTGTGCTGTGCTGTGCTG 806
 QY 161 LeuArgSerPheSerValLeuSerLeuAlaTrpGlyPheIleAlaAspValAspLeuGlu 180
 Db 807 CTGGGCTCTATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
 QY 181 SerAspLysTyrArgArgLeuGlnGlyLysMetArgPheThrLeuGlnGlyThrPheLeuArgLeu 200
 Db 867 AGTGAGAAAGTATCGCCCTGCGCCCTGCGGAGATGCCCTTCACTGCGGACCTTCCTGCT 926
 QY 201 AlaAlaLeuArgThrTyrArgGlyArgLeuAlaThrLeuProValGlyArgValGlyPhe 220
 Db 927 GCAAGCCTCGCATCTACCAAGGCCAAGCTGCTTCTCTGTAGAGACGTGGGCTCT 986
 QY 221 LysThrProAlaSerProValValValGlnGlnGlyProValAspAlaHisLeuValPro 240
 Db 987 AAGAGACCGGCTCT--ACACGTGTGCAGAAAGGGCCCGTGCACACACACTTGTTCCT 1043
 QY 241 LeuGlnGlnGlnValProSerHisTyrPglValValProAspGluAspPheValLeuVal 260
 Db 1044 CTGAGAGAGACAGCTCTTCTCATTTGACATGAGCTGTGACCAAGACAGACTTCGCTGGG 1103
 QY 261 LeuAlaLeuHisSerHisLeuAlaSerGlnMetPheAlaAlaProMetGlyArgCys 280
 Db 1104 CTGCTGTCTACACACCACTGAGCTCGAGCTGTGGAGCAACCATGGGCGCTGT 1163

```

QY      281 AlaIaGIYValMetHisLeuPheTYrValArgAlaGIYValSerArgAlaMetLeu 300
Db      1164 GAGCTGTGTATGATCTGTCTACGTACGTGGGGGTGTCAAGGGCTGCGCTGTG 1223
QY      301 ArgLeuPheLeuAlaMetGIuLYSGLYArgHisMetGIuTYrGIuCYsProTYrLeuVal 320
Db      1224 CGCCTTCTCGGCATGCAGAAAGGCAGACATATGACATGTGACTGTCCATACCTGTT 1283
QY      321 TYrValProValValAlaPheArgLeuGIuProLYSAspGIuLYSGLYValPheAlaVal 340
Db      1284 CATGTGCCCGTGTCTTCCGCTGTGAGCCAGAGCCAGAGGGCGTGTCTGTG 1343
QY      341 AspGIuLeuMetValSerGIuAlaValGIuLYSGLYValHisProAsnTYrPheTrp 360
Db      1344 GATGAGAGCTGATGATGTGAAGCTGTGCAGGGCCAAAGTGCACCCAAACTACCTTTGG 1403
QY      361 MetValSerGIuCYsValGIuProProSerTrpLYsProGIuInMetProProPro 380
Db      1404 ATGCTGTGTGCACAGAGATGCCCATCCGGCGGGGACTCCGGGGGGGCCACTCCA 1463
QY      381 GIuGIuPro 383
Db      1464 GAAGAACCA 1472

```

Search completed: July 12, 2003, 19:15:25
 Job time : 3460 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 17:20:27 ; Search time 18 seconds

(without alignments)
884.828 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGGRCVGLPFCRCVLT.....CVEPPPSKPKQMPPEEPL 384

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1962	97.4	384	1	SPH1_HUMAN
2	874	43.4	617	1	SPH2_MOUSE
3	856.5	42.5	654	1	SPH2_HUMAN
4	106.5	5.3	433	1	Y036_SYNY3
5	102.5	5.1	1065	1	K0G1_HUMAN
6	96	4.8	712	1	DNLI_RHOMR
7	94.5	4.7	260	1	PPNK_CAUCR
8	94	4.6	1247	1	IRBP_HUMAN
9	93	4.6	297	1	BMRU_BACSV
10	93	4.6	425	1	BIOA_SERMA
11	92.5	4.6	439	1	YSM3_CAEL
12	91	4.5	533	1	SERA_RAT
13	90	4.5	533	1	SH7_MOUSE
14	90	4.5	548	1	SERA_HUMAN
15	89	4.4	734	1	PSAB_PORPU
16	88.5	4.4	392	1	RURE_ACICA
17	86.5	4.3	259	1	PPNK_XYTRA
18	86.5	4.3	791	1	K0GL_DROME
19	86	4.3	567	1	GPV_RAT
20	86	4.3	998	1	ECAS_ARATH
21	85.5	4.2	334	1	G3P_PYRAB
22	85	4.2	827	1	K0GL_CAEL
23	84.5	4.2	300	1	YCP8_BRAZA
24	83.5	4.1	508	1	GLPK_MICPN
25	83.5	4.1	808	1	DHG_GLUOX
26	83	4.1	758	1	CHLD_TOBAC
27	83	4.1	3011	1	POLG_HCVH
28	82.5	4.1	554	1	COXA_PARDE
29	82.5	4.1	572	1	POXB_ECOLI
30	82	4.1	334	1	DPFV_ECOLI
31	82	4.1	448	1	SH7_RAT
32	82	4.1	450	1	TBB_BOMO
33	82	4.1	608	1	HFAC_CAUCR

34	82	4.1	2672	1	GCN1_YEAST
35	81.5	4.0	320	1	YRIB_ECOLI
36	81.5	4.0	1286	1	IRBP_BOVIN
37	81.5	4.0	1400	1	RON_HUMAN
38	81.5	4.0	2470	1	NTC2_MOUSE
39	81	4.0	396	1	DXR_DEIRA
40	81	4.0	419	1	PMT4_TOBAC
41	81	4.0	782	1	SMAB_MOUSE
42	80.5	4.0	335	1	G3P2_METAC
43	80.5	4.0	737	1	G3P_METTH
44	80.5	4.0	335	1	ETRI_BRAOL
45	80	4.0	319	1	COA2_POVMA

ALIGNMENTS

RESULT 1	SPH1_HUMAN	STANDARD:	PRT:	384 AA.
ID	Q9NYL1: Q9NYL13, Q9HD92, Q9NY70;			
AC	Q9NYL1: Q9NYL13, Q9HD92, Q9NY70;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Sphingosine kinase 1 (EC 2.7.1.-) (SK 1) (SPK 1).			
GN	SPHK1 OR SPHK OR SPK.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxId:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-20323213; PubMed-10863092;			
RA	Wetender A.J., Carlos-Dias E., Gossink M., Allen J.M., Takacs L.;			
RT	"Human sphingosine kinase: molecular cloning, functional			
RL	characterization and tissue distribution.";			
RL	gene 251:19-26(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX	MEDLINE-20263733; PubMed-10802064;			
RA	Nava V.E., Lacana E., Poultou S., Liu H., Suglura M., Kono K.,			
RT	Malsten S., Kohama T., Spiegel S.;			
RL	"Functional characterization of human sphingosine kinase-1.";			
RN	[3]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX	MEDLINE-20407120; PubMed-10947957;			
RA	Pitson S.M., D'Andrea R.J., Vandeleur L., Moretti P.A.B., Xia P.,			
RT	Gamble J.R., Vadas M.A., Wattenberg B.W.;			
RL	"Human sphingosine kinase: purification, molecular cloning and			
RL	characterization of the native and recombinant enzymes.";			
RL	Biochem. J. 350:429-441(2000).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Van Veldhoven P.P., Gijbers S.;			
RT	Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Ovary, and Mammary gland;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Magatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yanamoto K., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,			
RT	Nimomiya K., Iwayanagi T.;			
RL	"NEO human cDNA sequencing project.";			
CC	Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.			
CC	-I- FUNCTION: Catalyzes the phosphorylation of sphingosine to form			
CC	sphingosine 1-phosphate (SPP), a lipid mediator with both intra-			
CC	and extracellular functions. Also acts on D-erythro-sphingosine			
CC	and to a lesser extent sphinganine, but not other lipids, such as			
CC	D,L-threo-dihydrosphingosine, N,N-dimethylsphingosine.			

May 00

Sep 00

Jun 00

diacylglycerol, ceramide, or phosphatidylinositol.
 -1- CATALYTIC ACTIVITY: sphingosine + ATP -> sphingosine 1-phosphate + ADP.
 -1- SUBUNIT: Blinds to calmodulin.
 -1- TISSUE SPECIFICITY: Widely expressed with highest levels in adult liver, kidney, heart and skeletal muscle.
 -1- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL; AF26756; AAF73470.1; -
 CC EMBL; AF238083; AAF73423.1; -
 CC EMBL; AF200328; AAG01980.1; -
 CC EMBL; AK023393; BAB14558.1; -
 CC EMBL; AK022402; BAB14028.1; -
 CC EMBL; AJ245504; CAB92131.1; -
 CC Genbank; HGNC:11240; SPHK1.
 CC MIM; 603730; -
 CC InterPro; IPR001206; DAGKC.
 CC InterPro; IPR003622; DAG_kin_cat.
 CC Pfam; PF00781; DAGKC; 2.
 CC ProDom; PD005043; DAG_kin_cat; 1.
 CC SMART; SM00046; DAGKC; 1.
 CC Transferase; Kinase; ATP-binding; Calmodulin-binding.
 CC FT CONFILCT 6 MISSING (IN REF. 4).
 CC FT CONFILCT 11 15 LPRPC -> ARU (IN REF. 4).
 CC FT CONFILCT 114 115 NA -> KP (IN REF. 4).
 CC FT CONFILCT 251 251 V -> M (IN REF. 2).
 CC FT CONFILCT 260 260 V -> I (IN REF. 2).
 CC FT CONFILCT 302 302 L -> F (IN REF. 2).
 CC FT CONFILCT 325 325 V -> G (IN REF. 4).
 CC FT CONFILCT 337 337 V -> M (IN REF. 3).
 CC FT CONFILCT 384 AA; 42517 MW; EB04AF2034C2DB0 CRC64;
 CC SEQ SEQUENCE
 Query Match 97.4%; Score 1962; DB 1; Length 384;
 Best Local Similarity 97.7%; Pred. No. 2.9e-164;
 Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

RESULT 2
 SPH2_MOUSE STANDARD; PRT; 617 AA.
 ID SPH2_MOUSE
 AC Q9JIA7; Q9DBH6; Q91VA9;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sphingosine kinase 2 (EC 2.7.1.1-) (SK 2) (SPK 2).
 GN SPHK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC STRAIN=ICR; TISSUE=Brain;
 RX MEDLINE=20347850; PubMed=10751414;
 RA Liu H., Sugitara M., Nawa V.E., Edsall L.C., Kono K., Poulton S., Mlatien S., Kohama T., Spiegel S.;
 RT "Molecular cloning and functional characterization of a novel mammalian sphingosine kinase type 2 isoform."
 RL J. Biol. Chem. 275:19513-19520(2000).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Thompson D., Pyne S.;
 RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fieschmann W., Gaasterland T., Glass C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schirral L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G., Bromstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriquez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Segal T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohitsuki S., Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the phosphorylation of sphingosine to form sphingosine 1-phosphate (SP1), a lipid mediator with both intra- and extracellular functions. Also acts on D-erythro-dihydrosphingosine, D-erythro-sphingosine and L-threo-dihydrosphingosine.
 CC -1- CATALYTIC ACTIVITY: sphingosine + ATP -> sphingosine 1-phosphate + ADP.
 CC -1- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AF245448; AAF74125.1; -

DR EMBL: AF415214; AAL07500.1; -
 DR EMBL: AK004951; BAB23694.1; -
 DR EMBL: BC006941; AAH06941.1; -
 DR MGD: MGI:1861380; SPHK2.
 DR InterPro: IPR001206; DAGKC.
 DR InterPro: IPR003622; DAG_Kin_cat.
 DR Pfam: PF00781; DAGKC. 1.
 DR ProDom: PD005043; DAG_Kin_cat; 1.
 DR SMART: SM00046; DAGKC. 1.
 DR Transferase; Kinase; ATP-binding.
 FT CONFLICT 252 252 N -> S (IN REF. 1).
 FT CONFLICT 510 510 P -> T (IN REF. 1).
 FT CONFLICT 548 548 L -> F (IN REF. 1).
 FT CONFLICT 548 548 L -> F (IN REF. 1).
 SO SEQUENCE 617 AA; 65618 MW; 40EE2C2C288BE26A CMC64;

Query Match 43.4%; Score 874; DB 1; Length 617;
 Best Local Similarity 38.9%; Pred. No. 9.6e-69;
 Matches 187; Conservative 62; Mismatches 112; Indels 120; Gaps 4;

OY 10 VLPPRCVLYLNPGRGKALOLFRRSHVOPFLAEAEISFTLMTERRNHARELYSEEL 69
 DB 141 LPPKPRLLLYVNPFGGRLGALMQRKMDHVPVISEAGLSFTLQTERONHARELYVQSL 200
 OY 70 GRMDALVMSGDGLMHEVNGIMERPDMETAIQKPLCSIPAGSGNALAASINHYAGYEOV 129
 DB 201 SEMEGIVTVSGDGLLYEVNLGLDRPDMEDAVRMPIGVLPCCSGNALAGAVNHGGEQV 260
 OY 130 TNEDLLNCTLLLCRPVLSPPNLLSLTASGLRSFVSLAMGFADVLESDYRRIGE 189
 DB 261 VGVLLNCSLLRCRGSGHPDLISVTLASGRCFSFSLVAMGFLSDVDIHSEFRALGS 320
 OY 190 MRFTLGFRLAATRYRGLATLPV----- 215
 DB 321 ARFTLGAVLGLASLHTTRGRSLVPAETEPALPIGHSLSPPAKSELVAPADAPAAHSP 380
 OY 216 ----- 215
 DB 381 LHRVSDELPLPQPALVSPSPPEPLDLSLNGGGPELTGDMGAGDAPLSPDLLPSSP 440
 OY 216 -----GVRGKTP--ASPVVVOGCPVDAHLVPLEEQVPSHMQVVD 254
 DB 441 NALTAQLSPVIAEGPPEMPASSGFLPPIHSAPEASTWGPVHLLPLPGLPDPDWTEG 500
 OY 255 EDFVLVIALHSHLASEMFAPMGRCAAGVHLYFVRAGVSRAMLRLFLAMERGRHEX 314
 DB 501 E-FLVMGILPSHLCADLMAAPRAHAFDDGVHLCWVSGISRALLRLTLMERGNHPSL 559
 OY 315 ECPVLYVVPVAFRLERPKGKGFVAVDGLMNVSEAVOGVHPNRYFMVSGCCEPPSPWP 374
 DB 560 GCPHLYGAARAFAFLERPLPRGLLTVSGELVEYGPIDQAVRPLATLTG----PAGCKP 615
 OY 375 Q 375
 DB 616 Q 616

RESULT 3
 SPH2_HUMAN STANDARD; PRT; 654 AA.
 ID SPH2_HUMAN
 AC Q9NRA0; Q9NRA07; Q9H0Q2; Q9BRN1;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sphingosine kinase 2 (EC 2.7.1.-) (SK 2) (SPK 2).
 GN SPHK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
 RX MEDLINE-20347850; PubMed-10751414;
 RA Liu H., Sugitara M., Nava V.E., Edsall L.C., Kono K., Poulton S.,

RA Milstien S., Kohama T., Spiegel S.;
 RT "Molecular cloning and functional characterization of a novel
 RT mammalian sphingosine kinase type 2 isoform.";
 RT J. Biol. Chem. 275:19513-19520(2000).
 RN [2].
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Brain;
 RX MEDLINE-21154917; PubMed-11230166;
 RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansorge W., Boecker M., Bloecher H., Bauerbach S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehner K., Strack N.,
 RA Mewes H.-W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wandt R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RT Genome Res. 11:422-435(2001).
 RN [3].
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE-Eye, and Lymph;
 RA Strausberg R.;
 RP Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [4].
 RP SEQUENCE OF 1-354 FROM N.A. (ISOFORM 3).
 RC TISSUE-Carcinoma;
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isigai T., Sugano S.;
 RT "NDO human cDNA sequencing project.";
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Catalyzes the phosphorylation of sphingosine to form
 CC sphingosine 1-phosphate (SPP), a lipid mediator with both intra-
 CC and extracellular functions. Also acts on D-erythro-
 CC dihydrosphingosine, D-erythro-sphingosine and L-threo-
 CC dihydrosphingosine.
 CC -1- CATALYTIC ACTIVITY: Sphingosine + ATP - sphingosine 1-phosphate +
 CC ADP.
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; may be
 CC produced by alternative splicing
 CC -1- SIMILARITY: BELONGS TO THE SPHINGOSINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL: AF245447; AAF74124.1; -
 DR EMBL: AL136701; CAB66636.1; -
 DR EMBL: BC006161; AAH06161.1; -
 DR EMBL: BC010671; AAH10671.1; -
 DR EMBL: AK000599; BAA91280.1; -
 DR MIM: 607092; -
 DR Genew: HGNC:18859; SPHK2.
 DR InterPro: IPR001206; DAGKC.
 DR InterPro: IPR003622; DAG_Kin_cat.
 DR Pfam: PF00781; DAGKC. 1.
 DR ProDom: PD005043; DAG_Kin_cat; 1.
 DR SMART: SM00046; DAGKC. 1.
 KW Transferase; Kinase; ATP-binding; Alternative splicing.
 FT TRANSLOC 1 36 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 FT VARSPLIC 292 390
 FT LSVAMGEVSDVSDISERFRLGASAFRLGVLSTVLASGRFESF
 FT GRLSYLPATVEPASPTP -> PREDSDSSTSSACPLMTTTR
 FT RSCPRRAASMPGCPILPQALGFSRPIODRVNCGGGRIG
 FT SLTCRGTGRTLPAPARGGGSFLKLNINIVICKKKKK
 FT (IN ISOFORM 3).
 FT CONFLICT 49 49 P -> S (IN REF. 2).
 FT SEQUENCE 654 AA; 69217 MW; F73FCEC930DA50P CMC64;
 SO SEQUENCE 654 AA; 69217 MW; F73FCEC930DA50P CMC64;
 Query Match 42.5%; Score 856.5; DB 1; Length 654;
 Best Local Similarity 38.6%; Pred. No. 3.5e-67;


```

CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHOSBO-ESTER AND DAG
CC BINDING DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 1 MARCKS HOMOLOG REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AF061936; AAC62010.1; -
CC DR EMBL; AF219939; AAF43006.1; -
CC DR EMBL; AF219907; AAF43006.1; JOINED.
CC DR EMBL; AF219908; AAF43006.1; JOINED.
CC DR EMBL; AF219909; AAF43006.1; JOINED.
CC DR EMBL; AF219910; AAF43006.1; JOINED.
CC DR EMBL; AF219911; AAF43006.1; JOINED.
CC DR EMBL; AF219912; AAF43006.1; JOINED.
CC DR EMBL; AF219913; AAF43006.1; JOINED.
CC DR EMBL; AF219914; AAF43006.1; JOINED.
CC DR EMBL; AF219915; AAF43006.1; JOINED.
CC DR EMBL; AF219916; AAF43006.1; JOINED.
CC DR EMBL; AF219917; AAF43006.1; JOINED.
CC DR EMBL; AF219918; AAF43006.1; JOINED.
CC DR EMBL; AF219919; AAF43006.1; JOINED.
CC DR EMBL; AF219920; AAF43006.1; JOINED.
CC DR EMBL; AF219921; AAF43006.1; JOINED.
CC DR EMBL; AF219922; AAF43006.1; JOINED.
CC DR EMBL; AF219923; AAF43006.1; JOINED.
CC DR EMBL; AF219924; AAF43006.1; JOINED.
CC DR EMBL; AF219925; AAF43006.1; JOINED.
CC DR EMBL; AF219926; AAF43006.1; JOINED.
CC DR EMBL; AF219927; AAF43006.1; JOINED.
CC DR EMBL; AF219928; AAF43006.1; JOINED.
CC DR EMBL; AF219929; AAF43006.1; JOINED.
CC DR EMBL; AF219930; AAF43006.1; JOINED.
CC DR EMBL; AF219931; AAF43006.1; JOINED.
CC DR EMBL; AF219932; AAF43006.1; JOINED.
CC DR EMBL; AF219933; AAF43006.1; JOINED.
CC DR EMBL; AF219934; AAF43006.1; JOINED.
CC DR EMBL; AF219935; AAF43006.1; JOINED.
CC DR EMBL; AF219936; AAF43006.1; JOINED.
CC DR EMBL; AF219937; AAF43006.1; JOINED.
CC DR EMBL; AF219938; AAF43006.1; JOINED.
CC DR EMBL; AF219939; AAF43006.1; JOINED.
CC DR EMBL; AF219940; AAF43006.1; JOINED.
CC DR EMBL; AF219941; AAF43006.1; JOINED.
CC DR EMBL; AF219942; AAF43006.1; JOINED.
CC DR EMBL; AF219943; AAF43006.1; JOINED.
CC DR EMBL; AF219944; AAF43006.1; JOINED.
CC DR EMBL; AF219945; AAF43006.1; JOINED.
CC DR EMBL; AF219946; AAF43006.1; JOINED.
CC DR EMBL; AF219947; AAF43006.1; JOINED.
CC DR EMBL; AF219948; AAF43006.1; JOINED.
CC DR EMBL; AF219949; AAF43006.1; JOINED.
CC DR EMBL; AF219950; AAF43006.1; JOINED.
CC DR EMBL; AF219951; AAF43006.1; JOINED.
CC DR EMBL; AF219952; AAF43006.1; JOINED.
CC DR EMBL; AF219953; AAF43006.1; JOINED.
CC DR EMBL; AF219954; AAF43006.1; JOINED.
CC DR EMBL; AF219955; AAF43006.1; JOINED.
CC DR EMBL; AF219956; AAF43006.1; JOINED.
CC DR EMBL; AF219957; AAF43006.1; JOINED.
CC DR EMBL; AF219958; AAF43006.1; JOINED.
CC DR EMBL; AF219959; AAF43006.1; JOINED.
CC DR EMBL; AF219960; AAF43006.1; JOINED.
CC DR EMBL; AF219961; AAF43006.1; JOINED.
CC DR EMBL; AF219962; AAF43006.1; JOINED.
CC DR EMBL; AF219963; AAF43006.1; JOINED.
CC DR EMBL; AF219964; AAF43006.1; JOINED.
CC DR EMBL; AF219965; AAF43006.1; JOINED.
CC DR EMBL; AF219966; AAF43006.1; JOINED.
CC DR EMBL; AF219967; AAF43006.1; JOINED.
CC DR EMBL; AF219968; AAF43006.1; JOINED.
CC DR EMBL; AF219969; AAF43006.1; JOINED.
CC DR EMBL; AF219970; AAF43006.1; JOINED.
CC DR EMBL; AF219971; AAF43006.1; JOINED.
CC DR EMBL; AF219972; AAF43006.1; JOINED.
CC DR EMBL; AF219973; AAF43006.1; JOINED.
CC DR EMBL; AF219974; AAF43006.1; JOINED.
CC DR EMBL; AF219975; AAF43006.1; JOINED.
CC DR EMBL; AF219976; AAF43006.1; JOINED.
CC DR EMBL; AF219977; AAF43006.1; JOINED.
CC DR EMBL; AF219978; AAF43006.1; JOINED.
CC DR EMBL; AF219979; AAF43006.1; JOINED.
CC DR EMBL; AF219980; AAF43006.1; JOINED.
CC DR EMBL; AF219981; AAF43006.1; JOINED.
CC DR EMBL; AF219982; AAF43006.1; JOINED.
CC DR EMBL; AF219983; AAF43006.1; JOINED.
CC DR EMBL; AF219984; AAF43006.1; JOINED.
CC DR EMBL; AF219985; AAF43006.1; JOINED.
CC DR EMBL; AF219986; AAF43006.1; JOINED.
CC DR EMBL; AF219987; AAF43006.1; JOINED.
CC DR EMBL; AF219988; AAF43006.1; JOINED.
CC DR EMBL; AF219989; AAF43006.1; JOINED.
CC DR EMBL; AF219990; AAF43006.1; JOINED.
CC DR EMBL; AF219991; AAF43006.1; JOINED.
CC DR EMBL; AF219992; AAF43006.1; JOINED.
CC DR EMBL; AF219993; AAF43006.1; JOINED.
CC DR EMBL; AF219994; AAF43006.1; JOINED.
CC DR EMBL; AF219995; AAF43006.1; JOINED.
CC DR EMBL; AF219996; AAF43006.1; JOINED.
CC DR EMBL; AF219997; AAF43006.1; JOINED.
CC DR EMBL; AF219998; AAF43006.1; JOINED.
CC DR EMBL; AF219999; AAF43006.1; JOINED.
CC DR EMBL; AF220000; AAF43006.1; JOINED.
CC DR EMBL; AF220001; AAF43006.1; JOINED.
CC DR EMBL; AF220002; AAF43006.1; JOINED.
CC DR EMBL; AF220003; AAF43006.1; JOINED.
CC DR EMBL; AF220004; AAF43006.1; JOINED.
CC DR EMBL; AF220005; AAF43006.1; JOINED.
CC DR EMBL; AF220006; AAF43006.1; JOINED.
CC DR EMBL; AF220007; AAF43006.1; JOINED.
CC DR EMBL; AF220008; AAF43006.1; JOINED.
CC DR EMBL; AF220009; AAF43006.1; JOINED.
CC DR EMBL; AF220010; AAF43006.1; JOINED.
CC DR EMBL; AF220011; AAF43006.1; JOINED.
CC DR EMBL; AF220012; AAF43006.1; JOINED.
CC DR EMBL; AF220013; AAF43006.1; JOINED.
CC DR EMBL; AF220014; AAF43006.1; JOINED.
CC DR EMBL; AF220015; AAF43006.1; JOINED.
CC DR EMBL; AF220016; AAF43006.1; JOINED.
CC DR EMBL; AF220017; AAF43006.1; JOINED.
CC DR EMBL; AF220018; AAF43006.1; JOINED.
CC DR EMBL; AF220019; AAF43006.1; JOINED.
CC DR EMBL; AF220020; AAF43006.1; JOINED.
CC DR EMBL; AF220021; AAF43006.1; JOINED.
CC DR EMBL; AF220022; AAF43006.1; JOINED.
CC DR EMBL; AF220023; AAF43006.1; JOINED.
CC DR EMBL; AF220024; AAF43006.1; JOINED.
CC DR EMBL; AF220025; AAF43006.1; JOINED.
CC DR EMBL; AF220026; AAF43006.1; JOINED.
CC DR EMBL; AF220027; AAF43006.1; JOINED.
CC DR EMBL; AF220028; AAF43006.1; JOINED.
CC DR EMBL; AF220029; AAF43006.1; JOINED.
CC DR EMBL; AF220030; AAF43006.1; JOINED.
CC DR EMBL; AF220031; AAF43006.1; JOINED.
CC DR EMBL; AF220032; AAF43006.1; JOINED.
CC DR EMBL; AF220033; AAF43006.1; JOINED.
CC DR EMBL; AF220034; AAF43006.1; JOINED.
CC DR EMBL; AF220035; AAF43006.1; JOINED.
CC DR EMBL; AF220036; AAF43006.1; JOINED.
CC DR EMBL; AF220037; AAF43006.1; JOINED.
CC DR EMBL; AF220038; AAF43006.1; JOINED.
CC DR EMBL; AF220039; AAF43006.1; JOINED.
CC DR EMBL; AF220040; AAF43006.1; JOINED.
CC DR EMBL; AF220041; AAF43006.1; JOINED.
CC DR EMBL; AF220042; AAF43006.1; JOINED.
CC DR EMBL
```

```

FT DOMAIN 526 683 CATALYTIC-B (POTENTIAL) .
FT REPEAT 958 990 ANK 1.
FT REPEAT 997 1026 ANK 2.
FT DOMAIN 20 31 POLY-ALA.
FT DOMAIN 69 74 POLY-SER.
FT DOMAIN 95 102 POLY-ALA.
FT VARIANT 153 153 L -> F.
FT CONFLICT 160 160 /FTID=VAR_010190.
FT SEQUENCE 1065 AA; 116996 MW; B84971AA7630A799 CRC64;

Query Match 5.1%; Score 102.5; DB 1; Length 1065;
Best Local Similarity 24.9%; Pred. No. 0.66;
Matches 57; Conservative 30; Mismatches 61; Indels 81; Gaps 12;

OY 3 PAGEPRGVLPAPCRVLYLVNPRGKGG--KALQIFRSVOPLLAEAISFTLMTERRNH 59
DB 368 PISPP--LMP--LIVFVNPKSGSGNGCTKVLQFPMFTLP----- 403
OY 60 ARELVSEELGRMDAL-----VMSGDGLMEHVNGVLMERPDMETAIQPLCSL 108
DB 404 -ROYFDLSQEGPKDALELYRKVPNLRILACGDSGTGVIILIDE--LQLSQPPVGL 459
OY 109 PAGSNMAALASLNRYAGTEQYTNEEDLTNCLLLCR-----PVLSF 149
DB 460 PLGTGNDLARIILNMGGY--TDEPV--SKILQVEDGTGVOLDRMNHLVERNPDLPP 512
OY 150 MNL-----LSLHTASGLRSPFSYSLAWGFADYDLESDKRYRGLGEMPF 192
DB 513 EELEDGVCKLPLNFVNNY--FSL-----GPDHATYLEHERENPEKF 554

RESULT 6
DNLJ_RHOMR
ID DNLJ_RHOMR STANDARD: PRT: 712 AA.
AC P49421.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD+]).
GN LIGA OR LIG.
OS Rhodothermus marinus (Rhodothermus obamensis).
OC Bacteria; Bacteroidetes; Sphingobacteria; Sphingobacteriales;
OC Crenotrichiaceae; Rhodothermus.
OX NCBI_TaxID=29549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R-21;
RX MEDLINE=95369716; PubMed=7642120;
RA Thoburnardocott S.H., Jonsson Z.O., Andersson O.S.,
RA Kristjansson J.K., Eggertsson G., Palsson A.;
RT "Cloning and sequence analysis of the DNA ligase-encoding gene of
RT Rhodothermus marinus, and overproduction, purification and
RT characterization of two thermophilic DNA ligases.";
RL Gene 161:1-6(1995).
CC -!- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER
CC LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-
CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR
CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF
CC DAMAGED DNA. THIS ENZYME IS THERMOSTABLE BEING ACTIVE AT 5-75
CC DEGREES CELSIUS WITH APPARENT OPTIMAL ACTIVITY ABOVE 55 DEGREES
CC CELSIUS.
CC -!- CATALYTIC ACTIVITY: NAD(+) + [deoxyribonucleotide](N) +
CC [deoxyribonucleotide](M) -> AMP + nicotinamide nucleotide +
CC [deoxyribonucleotide](N+M).
CC -!- MISCELLANEOUS: THIS ENZYME IS THERMOSTABLE.
CC -!- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BCTD DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way

```



```

CC      AND THE REVERSED SEQUENCE IS GIVEN IN REF.1, REF.2 AND REF.3.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isdb-sib.ch/announce/
CC      or send an email to license@isb.sib.ch).
CC      -----
DR      EMBL; M33875; AAA59453.1; -
DR      EMBL; M33864; AAA59453.1; JOINED.
DR      EMBL; M33865; AAA59453.1; JOINED.
DR      EMBL; M33866; AAA59453.1; JOINED.
DR      EMBL; M22453; AAA36126.1; -
DR      EMBL; J05253; AAC18875.1; ALT_INIT.
DR      EMBL; J03912; AAA59188.1; -
DR      EMBL; X53044; CAA37213.1; -
DR      PIR; H24417; H24417.
DR      PIR; C25735; C25735.
DR      PIR; A31987; A31987.
DR      PIR; J50307; J50307.
DR      PIR; A33812; A33812.
DR      Genew; HGNC:9921; RBP3.
DR      MITM; 180290; -.
DR      InterPro; IPR003581; TSPC.
DR      Pfam; PF02692; IRBP; 4.
DR      SMART; SM00245; TSPc; 4.
DM      Vitamin A; Transport; Repeat; Signal; Glycoprotein.
FT      SIGNAL       1      17
FT      CHAIN        18     1247
FT                                     INTERPHOTORECEPTOR RETINOID-BINDING
FT                                     PROTEIN.
FT      DOMAIN       18     1230
FT      REPEAT       18     320
FT                                     1. 4 X APPROXIMATE TANDEM REPEATS.
FT      REPEAT       321    630
FT                                     2.
FT      REPEAT       631    931
FT                                     3.
FT      REPEAT       932    1230
FT                                     4.
FT      CARBOHYD     205     205
FT                                     N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD     515     515
FT                                     N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE     1247 AA; 135362 MW; 6C1841411E012E0F CRC64;

Query Match          4.7%; Score 94; DB 1; Length 1247;
Best Local Similarity 25.6%; Pred. No. 4.6;
Matches 79; Conservative 26; Mismatches 105; Indels 98; Gaps 19;

QY      PAGSGNLAASLNATVAGYEQ-----VNEDLLTNCT-----LLCRPVLSPMN---LLS 154
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      482 PGGSSSAVPILLSTFGPEAGRVHLLFTTYDKRTNTITGHESFMELPGRYSTOGGVYLTT 541
QY      155 LH-TASGLRSESVL--SLANGFIADVDDLESIDKYRRLLGEKMRPTLTGFRLAALRTYRGRLA 211
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      542 SHRTATAAEPAFLMOSIGMATL-----VGRI--TAGNL-----LHRTVPLL 582
QY      212 TLPRGRGVGFKRPAPRVVVOGQRFVAHLVPLEEQYPSNH---QVPDEDFVLVALLHSHTL 268
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      583 DTPGGSLALTPLYTLTFIDNRSEA-----WLGGGVNPD--ATVLAEEALDK 625
QY      269 ASEK--FAAPYGRCAGVMHF---YYRAGV--SRAMLLRLFLEAMEKRHNHEYCOPVLVY 321
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      626 AQEVLERHQSIALVEGTHLEAHYARPEVVGQTALLRKLAGOAGR-----TAVD 678
QY      322 VPIVAFLR--EPKDGKG-----VFAVDGELMVSEAVOGOVHPNYFMNVSGCVPPPSKP 374
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      679 LESLASQLTADLQEVSDHRLRLVFHSPGELVVEA-----PPP---P 717
QY      375 QQPMPPEE 382
DB      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      718 PAVPSPEE 725

RESULT 9
ID      BMRU_BACSU STANDARD; PRF; 297 AA.
```

AC P39074;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein bmrV.
 GN BMRV.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / Marburg;
 RA MEDLINE=95050642; PubMed=7961792;
 RX Ahmed M., Borsch C.M., Taylor S.S., Vazquez-Laslop N., Neyfakh A.A.;
 RT "A protein that activates expression of a multidrug efflux
 transporter upon binding the transporter substrates."
 RL J. Biol. Chem. 269:28506-28513(1994).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RA MEDLINE=97124195; PubMed=8969508;
 RX Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
 RA Kobayashi Y.;
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
 the Bacillus subtilis genome containing the skin element and many
 sporulation genes."
 RL Microbiology 142:3103-3111(1996).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bartero M.G., Bessieres P., Bolotin A., Borchert S., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entlan K.D., Erington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karmata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudaga B., Park S.H.,
 RA Parro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,
 RA Plescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivoita C., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiguchi J., Sekowska E., Serot S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccotti E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Trosato V., Uchiyama S., Vandenberg M., Vannier F., Vassatoli A.,
 RA Viari A., Wambolt A., Wedler E., Wedler H., Weltzenegger T.,
 RA Winters P., Wipalt A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis."
 RL Nature 390:249-256(1997).
 CC -I- SIMILARITY: TO E.COLI YEGS AND TO SYNECHOCYSTIS PCC 6803 SLI0036.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L25604; AAB81538.1; -

DR EMBL; D84432; BAA12602.1; -
 DR EMBL; Z99116; CAB14331.1; -
 DR Subtilisin, BGI0302; bmrV.
 DR InterPro: IPR005218; Cons_hypoth147.
 DR InterPro: IPR001206; DAGKc.
 DR InterPro: IPR003622; DAG_Kin_cat.
 DR Pfam: PF00781; DAGKc; 1.
 DR ProDom: PD005043; DAG_Kin_cat; 1.
 DR SMART: SM00046; DAGKc; 1.
 DR TIGRFAMs: TIGR00147; TIGR00147; 1.
 KW COMPLETE PROTEOME.
 SQ SEQUENCE 297 AA; 32461 MW; 2EB5B9DC35F90CDD CRC64;
 Query Match 4.6%; Score 93; DB 1; Length 297;
 Best Local Similarity 19.1%; Pred. No. 0.87;
 Matches 68; Conservative 55; Mismatches 157; Indels 76; Gaps 13;
 QY 16 RVLVLNPRGGKRALQLFRSHVPLAEAFISTMLTERRNHARLVSEELGRDAL 75
 DB 4 RKALLIHGNANKNIKALKAVPVLSHILDEVYIKQTKKDAHYFCSID-DSYDLY 62
 QY 76 VMSGDGLMEVYVNL--MERPDWETAFORPLCS-LPAGSGNALASLNHYAGEQVYNE 132
 DB 63 FLIGDGTIHOCINAIISALR-----KAVGILPGGTNDSSRYLG----- 103
 QY 133 DLTNCTLLICRPVLSPNMLSLATASGLRSFVSLAMGFADVDESQYRRLGEMRF 192
 DB 104 -----IPQWLAK--AAEALMGKRTSV-----DYCCMDNRY----- 132
 QY 193 TLGTPLR-----LAALRYRGRRLATLPVGRVGEKTPSPVVOGQPVDAHLVLEQVP 246
 DB 133 ---FLNFWGIGLLETSSNQINETERKALIGKISTYSLRLRYSSAASPMPLKIDGEEI- 187
 QY 247 SHWQVVDDEFLVYLLSHLASEMFAPMGRCAAGVHMLFYRAGVSRMLRLFLAM 306
 DB 188 -----KEAWMLVMNQYIGTRNPLPDASIDGLDLYICRN--TNLTALRELSM 238
 QY 307 EKGRHEVECPVLYVYVAVFLERPKDGVAVANGELMVSAGVGVHNVFMVY 362
 DB 239 EOG-SIDRFTGELSVQASRIEITDTAKA-DMQGEVITRPVAVIYVLPQIHML 292
 RESULT 10
 BLOA_SERMA STANDARD; PRT; 425 AA.
 ID BLOA_SERMA
 AC P36568;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
 DE (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
 DE aminotransferase).
 GN BLOA.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Str41;
 RX MEDLINE=94071435; PubMed=8250549;
 RA Sakurai N., Imai Y., Masuda M., Komatsubara S., Tosa T.;
 RT "Molecular breeding of a biotin-hyperproducing Serratia marcescens
 strain."
 RL Appl. Environ. Microbiol. 59:3225-3232(1993).
 CC -I- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-
 CC oxononanoate = S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-
 CC diaminononanoate.
 CC -I- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -I- PATHWAY: Biotin biosynthesis.
 CC -I- SUBUNIT: HOMODIMER.
 CC -I- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D17468; BAA04284.1; -
DR HSSP: P12995; 1003.
DR InterPro: IPR000954; AminoTran_3.
DR Pfam: PF00202; aminoTran_3; 1.
DR TIGRFAMs: TIGR00508; bioA; 1.
DR PROSITE: PS00600; AA_TRANSFERR CLASS_3; 1.
DR Biotin biosynthesis; Transferrase; AminoTransferase;
KW Pyridoxal phosphate.
KW BINDING 277
FT SEQUENCE 425 AA; 4667 MM; 8BA657E59B894AF6 CRC64.
SQ
Query Match 4.6%; Score 93; DB 1; Length 425;
Best Local Similarity 20.3%; Pred. No. 1.4; Indels 106; Gaps 16;
Matches 71; Conservative 47; Mismatches 125;
QY 15 CRVIVLNPGRGKQALQ-LFRSHVQPLAEAEISFTLMTERRNHARELYRSEELG--- 70
DB 94 CRRLVEMTP-----EALQCVFLADSGSVAVEVSLKMAIQWQANGERRQRLTLRHGYHG 148
QY 71 -RMDALVYMSGDGLMHEVYNGLMERPDMETAIQKPLCSLPGSGNALAASLNHYAGYQV 129
DB 149 DTFAMSVCPDSDNSMHSLYQGYLAPHLFATA---POCRFD-----DEW 188
QY 130 TNEDLTNTCTLL-----ICRPVLSPMNLISLHTASGRSP-----SVLSAMFEI 174
DB 189 REEDIAPRALLEQHAGEVAIVLEPV-----VQAGGRITHTPYLKAVDVRVAIKLL 244
QY 175 ADVLESDDKRYRLGEM-----RFTLGTFRLIALRTYRGRLATLPVGRV 218
DB 245 LIADENATGFRGTGKLFCEHAQVVPDILCLGKALGTGVMYLSLTLTRHVAERTISNAA 304
QY 219 GFKTPASPVYVQGP-----VDAILVPLEEOVPSIMQVVPDEDFVLVIALLSH 267
DB 305 G-----CFMHGPTFMGNPLACAVDASLALLAE---NMWA-----QVSAIETQ 345
QY 268 LASEMFAPMGRG-----AAGVYHILFYRAGVSRAMLLFLFLAMEKG 309
DB 346 LKRELL-PLGRCRMPHVRVLAIGVEMREPVDAELQGEF-VERG 390
RESULT 11
YSM3_CAEEL
ID YSM3_CAEEL STANDARD: PRT; 439 AA.
AC Q10123.
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein F52C9.3 in chromosome III.
GN F52C9.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Favello T.;
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
RE REVIEWS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U39850; AAA81060.2; -
DR WormPep: F52C9.3; CE29806.
DR InterPro: IPR001206; DAGKC.
DR InterPro: IPR003622; DAG_Kin_cat.
DR PRODOM: PD005043; DAG_Kin_cat; 1.
DR SMART: SM00046; DAGKC; 1.
KW Hypothetical protein.
SQ
Query Match 4.6%; Score 92.5; DB 1; Length 439;
Best Local Similarity 26.5%; Pred. No. 1.6;
Matches 27; Conservative 14; Mismatches 58; Indels 3; Gaps 1;
QY 13 RCRVIVLNPGRGKQALQ-LFRSHVQPLAEAEISFTLMTERRNHARELYRSEELGRW 72
DB 66 RPKRFVLYVNEGSRGCFDQFNKALFLFLHAGVQDVADNOALEALGAVDQEA 125
QY 73 DALVYMSGDGLMHEVYNGLMERPDMETAIQKPLCSLPGSGN 114
DB 126 DILVYVGGDGTIGTVYIGIFRNRE---KAQLPVGRYPGGYDN 164
RESULT 12
SERV_RAT
ID SERV_RAT STANDARD: PRT; 533 AA.
AC 008651.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH).
GN PGDH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97250633; PubMed=9163325;
RA Achouri Y., Rider M.H., van Schaftingen E., Robbi M.;
RT "Cloning, sequencing and expression of rat liver 3-phosphoglycerate
RT dehydrogenase.";
RL Biochem. J. 323:365-370(1997).
CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) -> 3-
CC phosphohydroxypyruvate + NADH.
CC -1- PATHWAY: Serine biosynthesis; first step.
CC -1- SUBUNIT: Homotrimer.
CC -1- TISSUE SPECIFICITY: LIVER, KIDNEY, BRAIN, TESTIS.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X97772; CAA66374.1; -
DR HSSP: P08328; IPSD.
DR InterPro: IPR002162; D_2HydAc_dh.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH_C; 1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.

```

KM Serine biosynthesis: Oxidoreductase; NAD.
 FT ACT_SITE 236 SUBSTRATE-BINDING (BY SIMILARITY).
 FT ACT_SITE 265 BY SIMILARITY.
 FT ACT_SITE 283 BY SIMILARITY.
 SQ SEQUENCE 533 AA; 56493 MW; 7273DAC349E95EF CRC64;

Query Match
 Best Local Similarity 22.8%; Pred. No. 2.8;
 Matches 64; Conservative 45; Mismatches 96; Indels 76; Gaps 14;

51 LALTERNHARELYRSEELGRDALVYMSGDLMEHVNGL-----MERPDWE 98
 28 LQVEKONLSKEELIA-ELDDCEGLVRSATKVTADYINAEKLYVGRAGTGVNDVLE 86
 99 TAIGKPLCSLPAGSGNALASLNHYAGYEQVNTEDLLFNCTLLCRPVLSPMILSLHNA 158
 87 AATRGVLYMNTPNONSLSAA-----ELT-----CGMLCLARQIPQATASMK-- 129
 159 SGLRSFVLSLAMEFIADVLESDKYRRLG-EMR-FTLGTEFLIAL-RTYRGLATLPV 215
 130 -----DCKWDKKFMGTGLNGKTLG-ILGGRIGREVAARMOAFGM 169
 216 GHVGEKTPASPVY-----VQGPVDALVPLEEOPPSHMOVYPDEDEVLLALSHSLAS 270
 170 KTVGVDPLISPEVASFVQO-----LPLEE-----IWPICDFTVHTPLLPSTTG 215
 271 EHFAPMGRCAGVHLEFYVRAG-VSRAMLRLFLAMEKGR 310
 216 LINDSTFAOCKRGVRYVNCARGLYDEGALLR---ALDSQ 253

RESULT 13
 5H7_MOUSE STANDARD: PRT: 448 AA.
 ID 5H7_MOUSE
 AC P32304;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-1998 (Rel. 36, Last annotation update)
 DE 5-Hydroxytryptamine 7 receptor (5-HT-7) (5-HT-X) (serotonin receptor) (5HT7).
 GN HTR7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathia; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93360913; PubMed=8394987;
 RA Plassat J.-L., Amlaiky N., Hen R.;
 RT "Molecular cloning of a mammalian serotonin receptor that activates adenylyate cyclase."
 RL Mol. Pharmacol. 44:229-236(1993).
 CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATE ADENYLYATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z23107; CA80654.1; -
 DR PTR: S36402; S36402.
 DR MGD; MGI:99841; Htr7.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam; Pf00001; 7tm_1; 1.

DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR_FL_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECEPTOR_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; glycoprotein;
 KW Multigene family; Lipoprotein; Palmitate.
 FT DOMAIN 1 84
 FT TRANSMEM 85 111
 FT DOMAIN 112 121
 FT TRANSMEM 122 147
 FT DOMAIN 148 159
 FT TRANSMEM 160 181
 FT DOMAIN 182 201
 FT TRANSMEM 202 225
 FT DOMAIN 226 240
 FT TRANSMEM 241 263
 FT DOMAIN 264 331
 FT TRANSMEM 332 355
 FT DOMAIN 356 367
 FT TRANSMEM 368 390
 FT DOMAIN 391 448
 FT CARBOHYD 5 5
 FT CARBOHYD 69 69
 FT DISULFID 158 234
 FT LIPID 404
 SQ SEQUENCE 448 AA; 49861 MW; 01C7380300991C30 CRC64;

Query Match
 Best Local Similarity 23.7%; Pred. No. 2.7;
 Matches 78; Conservative 45; Mismatches 130; Indels 76; Gaps 21;

59 HARELYRSEELGRDALVYMSGDLMEHVNGLMER-----PDWETAIORPLCSLPAGSG 113
 15 HLRSLT-LPEVGR-RLDLSPDGGAHSVSSMPPHLSGPE-VTASPAFTMDAPRDNV 70
 114 NALAASLNHYAGYEQVNTEDLLFNCTLLI---CAPVLS-----PMNLISLHTRASG 160
 71 SCGGEQIN-YGREKVVIGSIITLLTLIAQNCIVASVCFVKNRQPSNYLVSILA-- 127
 161 LRSFVLSLAMEFIADVLESDKYRRLGEMRTLTGFLRL-----AALRTYRGLATLP 214
 128 LADLSVAVAVMPVSTDLIGRW-IFG--HFCVAFIAMDWCCTASIMT---LCVIS 180
 215 VGR-VGEKTPASPVVYVQGPVDALV-----PLEEOPV---SHW-QVYPDEDEVLYLAL 264
 181 IDRLGITRRLTYPRVONGCKAKMLISVPLSASTLPPLRGMAOVNDDVCLII----- 236
 265 HSHLASEMFAAPMGRCAGVHLEFYVRAGVSRAMLRLFLAMEK--GRHMEYECPLYVY 322
 237 -----SODFGTYTISTAVA-----FTIPMSVMLFMVQIYKARKSAKHESGFP----- 282
 323 PVVAFRLPEPKDKGVFAVDGLMVSEAVQ 351
 283 -----RVQP---ESVISLNGVVKLOKEVE 303

RESULT 14
 SERA_HUMAN STANDARD: PRT: 533 AA.
 ID SERA_HUMAN
 AC 043175; O9B001;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (3-PGDH).
 GN PGDH OR PGDH3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20179699; PubMed=10713460;
 RA Cho H.M., Jun D.Y., Bae M.A., Ahn J.D., Kim Y.H.;
 RT "Nucleotide sequence and differential expression of the human

3-phosphoglycerate dehydrogenase gene.";
 RL Gene 245:193-201(2000).
 [2]
 RP SEQUENCE FROM N.A., AND VARIANTS PHGDH DEFICIENCY MET-425 AND MET-490.
 RX MEDLINE-20530221; PubMed-1105895;
 RA Klomp L.W.J., de Koning T.J., Malingre H.E.M., van Beurden E.A.C.M.,
 RA Brink M., Opdam F.L., Duran M., Jaeken J., Plincha M.,
 RA van Maldergem L., Poll-The B.T., van den Berg I.E.T., Berger R.,
 RT "Molecular characterization of 3-phosphoglycerate dehydrogenase
 RT deficiency -- a neurometabolic disorder associated with reduced
 RT L-serine biosynthesis.";
 RL Am. J. Hum. Genet. 67:1389-1399(2000).
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, Lung, and Muscle;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) = 3-
 CC phosphohydroxypyruvate + NADH.
 CC -1 PATHWAY: Serine biosynthesis: first step.
 CC -1 SUBUNIT: Homotrimer (By similarity).
 CC -1 DISEASE: Defects in PHGDH are the cause of a deficiency
 CC characterized by congenital microcephaly, psychomotor retardation,
 CC and seizures.
 CC -1 SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 DR EMBL: AF006043; AAB88664.1; -
 DR EMBL: AF171237; AAD51415.1; -
 DR EMBL: BC000303; AAH00349.1; -
 DR EMBL: BC001349; AAH01349.1; -
 DR EMBL: BC011262; AAH11262.1; -
 DR HSSP: P08328; 1PSD.
 DR Gene: HGNC:8923; PHGDH.
 DR MIM: 606879; -
 DR MIM: 601815; -
 DR InterPro: IPR002162; D_2hyd-ac_dh.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH; 1.
 DR PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.
 KW Serine biosynthesis; Oxidoreductase; NAD; Disease mutation.
 FT ACT_SITE 236 236 SUBSTRATE-BINDING (BY SIMILARITY).
 FT ACT_SITE 265 265 BY SIMILARITY.
 FT ACT_SITE 283 283 BY SIMILARITY.
 FT VARIANT 425 425 V->M (IN PHGDH DEFICIENCY).
 FT VARIANT 490 490 V->M (IN PHGDH DEFICIENCY).
 FT VARIANT 490 490 V->M (IN PHGDH DEFICIENCY).
 FT CONFLICT 25 25 D->E (IN REF. 1).
 FT SEQUENCE 533 AA; 56650 MW; C58E872275C45B35 CRC64;
 SO
 Query Match 4.5%; Score 90; DB 1; Length 533;
 Best Local Similarity 19.6%; Pred. No. 3.4;
 Matches 61; Conservative 53; Mismatches 101; Indels 96; Gaps 14;
 OY 51 LMLERRHARELVESEELGRWDALVYVSGDLMEHVNGL-----MERPWE 98
 DB 28 LQVEKQNLSEKELLA-ELQDCEGLIVRSATKRYADVYNAEKLVYGRAGGVNDVLE 86
 OY 99 TAIQKPLCSLPAAGSNALASLNIHAYGEVYTNEDLTNCTLLCLCPVLSPNMLSLHTA 158
 DB 87 AATRGILVYMTYPNGNSLSAA-----ELT-----CGMIMCLAIQRIQ----- 123

159 SGLRSFVSLAMGFADVDLESDKYRRLGEMRFTL-GTFRLAL-----RTYGRRLATL 213
 DB 124 -----ATASKDKGKWEKKRNGELNKNKTIGILGRIGEVAFRNGSF 167
 OY 214 PVGRGKTPASPV-----VOQGVDAHLVPLEQVPSHWQVVP-----DEDF----- 257
 DB 168 GMRKIGYDPIISPEVSASFVQVQPLE-EIWPICDFIVHPPLPSTGTLNDWTFACK 226
 OY 258 --VIVALLHSHLASE--MFAAPMGRCAGVMHLFYRAVSAMLLRLFLAMEKGNM 312
 DB 227 KGVRYVNCARGIVDEGALLRALQSGCAGALVFTPEPRDRAL-----V 273
 OY 313 EYE-----CPYL 319
 DB 274 DHEVVISCPHL 284
 RESULT 15
 PSAB_PORPU STANDARD; PRT; 734 AA.
 AC P51285;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Photosystem I P700 chlorophyll A apoprotein A2 (PsaB) (PSI-B).
 CN PSAB.
 OS Porphyra purpurea.
 OG Chloroplast.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
 OX NCBI_TaxID=2787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Avonport;
 RA Reith M.E., Muniholland J.;
 RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
 RL genome.";
 RL Plant Mol. Biol. Rep. 13:333-335(1995).
 CC -1 FUNCTION: PsaA and psab bind P700, the primary electron donor of
 CC photosystem I (PSI), as well as the electron acceptors A0, A1, and
 CC FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin
 CC oxidoreductase.
 CC -1 COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1
 CC is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.
 CC -1 SUBUNIT: A psaa/B heterodimer binds the P700 chlorophyll special
 CC pair and subsequent electron acceptors. The PSI reaction center of
 CC higher plants and algae is composed of one at least 11 subunits.
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast
 CC thylakoid membrane.
 CC -1 SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

 DR EMBL: U38804; AAC08171.1; -
 DR HSSP: P25897; 1JUB0.
 DR InterPro: IPR001280; PSI_PsaA/B.
 DR Pfam: PF00223; psaa-psab; 1.
 DR PRINTS: PR00257; PROTSYPSAB.
 DR PROSITE: PS00419; PHOTOSYSTEM_I_PSAAB; 1.
 KW Photosynthesis; Photosystem I; Transports; Electron transport;
 KW Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;
 KW Iron-sulfur; 4Fe-4S; Chlorophyll.
 FT TRANSMEM 46 69 I (POTENTIAL).
 FT TRANSMEM 135 158 II (POTENTIAL).
 FT TRANSMEM 175 199 III (POTENTIAL).
 FT TRANSMEM 273 291 IV (POTENTIAL).
 FT TRANSMEM 330 353 V (POTENTIAL).
 FT TRANSMEM 369 395 VI (POTENTIAL).
 FT TRANSMEM

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 18:07:32 ; Search time 136 Seconds

(without alignments)
328.715 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015
Sequence: 1 MDPAGGPRGVLPKRCRVLV.....CPEPPSKPKQMPPEEPL 384

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 11641973 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTOS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2015	100.0	384	10 US-09-784-810A-2	Sequence 2, Appli
2	1962	97.4	384	10 US-09-970-516-2	Sequence 2, Appli
3	1954	97.0	384	10 US-09-796-487-3	Sequence 3, Appli
4	1594.5	79.1	382	10 US-09-784-810A-4	Sequence 4, Appli
5	1587.5	78.8	382	10 US-09-970-516-6	Sequence 4, Appli
6	1585.5	78.7	388	10 US-09-817-676A-15	Sequence 15, Appli
7	1585.5	78.7	388	10 US-09-796-487-2	Sequence 2, Appli
8	1582.5	78.5	381	10 US-09-796-487-1	Sequence 1, Appli
9	1582.5	78.5	381	10 US-09-796-487-4	Sequence 4, Appli
10	1564.5	77.6	373	10 US-09-796-487-5	Sequence 5, Appli
11	1467	73.8	293	9 US-10-023-282-328	Sequence 328, App
12	930.5	46.2	204	10 US-09-796-487-9	Sequence 9, Appli
13	875	43.4	617	10 US-09-817-676A-12	Sequence 12, Appli
14	856.5	42.5	618	10 US-09-817-676A-14	Sequence 14, Appli
15	856.5	42.5	618	10 US-09-970-516-4	Sequence 4, Appli
16	809.5	40.2	638	9 US-09-969-896-3	Sequence 3, Appli
17	453	22.5	119	9 US-10-023-282-788	Sequence 788, App
18	372.5	18.5	415	10 US-09-784-810A-15	Sequence 15, Appli
19	372	18.5	392	10 US-09-796-487-6	Sequence 6, Appli

20	369	18.3	380	10 US-09-784-810A-14	Sequence 14, Appli
21	355.5	17.6	424	10 US-09-796-487-8	Sequence 8, Appli
22	355.5	17.6	453	10 US-09-784-810A-12	Sequence 12, Appli
23	337	16.7	403	10 US-09-796-487-7	Sequence 7, Appli
24	263.5	13.1	436	10 US-09-784-810A-13	Sequence 13, Appli
25	263.5	13.1	537	9 US-09-969-896-10	Sequence 10, Appli
26	263.5	13.1	562	9 US-09-969-896-11	Sequence 11, Appli
27	250	12.4	471	10 US-09-784-810A-6	Sequence 6, Appli
28	223	11.1	326	9 US-09-969-896-2	Sequence 2, Appli
29	223	11.1	326	10 US-09-784-810A-11	Sequence 11, Appli
30	159	7.9	182	10 US-09-784-810A-29	Sequence 29, Appli
31	151	7.5	365	9 US-09-738-626-6336	Sequence 6336, Ap
32	147	7.3	421	9 US-09-948-820-50	Sequence 50, Appli
33	145	7.2	240	10 US-09-796-487-10	Sequence 10, Appli
34	137	6.8	319	9 US-10-156-761-9811	Sequence 9811, Ap
35	124	6.2	322	9 US-10-156-761-10532	Sequence 10532, A
36	113	5.6	296	10 US-09-815-242-4883	Sequence 4883, Ap
37	113	5.6	362	10 US-09-815-242-10793	Sequence 10793, A
38	111	5.5	294	10 US-09-815-242-13334	Sequence 13334, A
39	108	5.4	315	10 US-09-815-242-12862	Sequence 12862, A
40	102.5	5.1	1065	10 US-09-771-161A-239	Sequence 239, App
41	102	5.1	294	10 US-09-815-242-5505	Sequence 5505, Ap
42	102	5.1	315	10 US-09-815-242-12494	Sequence 12494, A
43	99	4.9	296	9 US-10-156-761-14222	Sequence 14222, A
44	94.5	4.7	2011	9 US-09-832-292-31	Sequence 31, Appli
45	93.5	4.6	511	9 US-10-156-761-10124	Sequence 10124, A

ALIGNMENTS

RESULT 1					
US-09-784-810A-2					
; Sequence 2, Application US/09784810A					
; Patent No. US20020082203A1					
; GENERAL INFORMATION:					
; APPLICANT: RASTELLI, LUCA					
; TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING					
; TITLE OF INVENTION: SAME					
; FILE REFERENCE: 10716-08					
; CURRENT APPLICATION NUMBER: US/09/784,810A					
; CURRENT FILING DATE: 2001-02-14					
; PRIOR APPLICATION NUMBER: 60/182,360					
; PRIOR FILING DATE: 2000-02-14					
; PRIOR APPLICATION NUMBER: 60/191,261					
; PRIOR FILING DATE: 2000-03-22					
; NUMBER OF SEQ ID NOS: 29					
; SOFTWARE: Patentin Ver. 2.1					
; SEQ ID NO 2					
; LENGTH: 384					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-09-784-810A-2					
Query Match					
Best local Similarity 100.0%; Pred. No. 3.4e-189;					
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MDPAGGPRGVLPKRCRVLV	LNPRGKGKALD	FRSHVPLAE	AFISFTLTERRNA 60
DB	1	MDPAGGPRGVLPKRCRVLV	LNPRGKGKALD	FRSHVPLAE	AFISFTLTERRNA 60
QY	61	RELVRSEELGRDVALVYNSG	DDGLMHEVYVNG	LMPERMPERAIQ	PICLSLPGSGNALAASL 120
DB	61	RELVRSEELGRDVALVYNSG	DDGLMHEVYVNG	LMPERMPERAIQ	PICLSLPGSGNALAASL 120
QY	121	NHYAGEQYTNEDLFTNCTLL	ICRPVLSPMNLS	LTASGLNSFVLS	LAMFIADVDLE 180
DB	121	NHYAGEQYTNEDLFTNCTLL	ICRPVLSPMNLS	LTASGLNSFVLS	LAMFIADVDLE 180
QY	181	SDKYRRLGEMRFTLGTFLRL	ALRLRYRGRLATL	PVGRVGF	KTPASPVVVOGQVDNAHLP 240
DB	181	SDKYRRLGEMRFTLGTFLRL	ALRLRYRGRLATL	PVGRVGF	KTPASPVVVOGQVDNAHLP 240


```

: GENERAL INFORMATION:
: APPLICANT: RASTELLI, LUCA
: TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
: TITLE OF INVENTION: SAME
: FILE REFERENCE: 10716-08
: CURRENT APPLICATION NUMBER: 05/09/784, 810A
: CURRENT FILING DATE: 2001-02-14
: PRIOR APPLICATION NUMBER: 60/182,360
: PRIOR FILING DATE: 2000-02-14
: PRIOR APPLICATION NUMBER: 60/191,261
: PRIOR FILING DATE: 2000-03-22
: NUMBER OF SEQ. ID NOS: 29
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4
: LENGTH: 362
: TYPE: PRT
: ORGANISM: Mus musculus
: IS-09-784-810A-4

```


Best Local Similarity 79.6%; Pred. No. 8.9e-147;
Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

```
QY 7 PRGLPPRCRVLLVNLNPGCGKALOLFRRSHVQPLAEAEISFTLMTERRNHARELYRS 66
Db 6 PRGLPPRCRVLLVNLNPGCGKALOLFRRSHVQPLAEAEISFTLMTERRNHARELYCA 65
QY 67 EELGRMALVMSGDGLMHEVNGIMERPDMETAIORPLCSLPAGSGNALAASLNHYAGY 126
Db 66 EELGRMALVMSGDGLMHEVNGIMERPDMETAIORPLCSLPAGSGNALAASLNHYAGY 125
QY 127 EQVTNEDLLNCTLLLCRPVLSMNLISLHTASGLRFSVLSLWGLFADVDESXKRR 186
Db 126 EQVTNEDLLNCTLLLCRRRLSPMNLISLHTASGLRFSVLSLWGLFADVDESXKRR 185
QY 187 LGEHRTLTGFLRLAALRTYGRGLATLPVGRVGFRTSPAPVVOGPPVDALVPLEBVP 246
Db 186 LGEHRTLTGFLRLAALRTYGRGLATLPVGRVGFRTSPAPVVOGPPVDALVPLEBVP 244
QY 247 SHMOVDEDEFLVLTALSHLASEMFAAPMGRCAGVMHLYYRAGYSRALLRLFLAM 306
Db 245 SHMTVPPEQDFLVLTALSHLASEMFAAPMGRCAGVMHLYYRAGYSRALLRLFLAM 304
QY 307 EKGHRMEYECPLYLVVPVAFRLPEPKDGKGFVAVDGLMSEAVOGVHPNYFMVSGCV 366
Db 305 OKGRHMEIDCPYLHVVPVAFRLPEPKDGKGFVAVDGLMSEAVOGVHPNYFMVSGCV 364
QY 367 EPPSKRQOMPPEEP 383
Db 365 DAPSGRDSRRGPPPEEP 381
```

RESULT 9

```
US-09-796-487-4
: Sequence 4, Application US/09796487
: Patent No. US20020042358A1
: GENERAL INFORMATION:
: APPLICANT: Spiegel, Sarah
: TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
: FILE REFERENCE: 07320001aa (2033957-0001)
: CURRENT APPLICATION NUMBER: US/09/796,487
: CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: US 60/186,532
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: US 09/530,868
: PRIOR FILING DATE: 2000-05-05
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 381
: TYPE: PRT
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (1)..(381)
: OTHER INFORMATION: SEQ ID NO 4 is the peptide sequence of mspk1a in Figure 3, corre
: OTHER INFORMATION: sponding to amino acid residue 124 to 504 of SPK1a of Genbank se
: OTHER INFORMATION: quence Accession Number AF068748. SEQ ID NO 4 is equivalent to SE
: OTHER INFORMATION: Q ID NO 1 that is the amino acid sequence of SPK1a in Figure 1.
: PUBLICATION INFORMATION:
: AUTHORS: Kohama et al.,
: TITLE: Molecular cloning and functional characterization of murine sphingosine
: JOURNAL: Journal of Biological Chemistry
: VOLUME: 273
: ISSUE: 37
: PAGES: 23722-23728
: DATE: 1998
: DATABASE ACCESSION NUMBER: AAC61697
: DATABASE ENTRY DATE: 1998-09-26
: RELEVANT RESIDUES: (124)..(504)
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: AAC61697
```

DATABASE ENTRY DATE: 1998-09-26
RELEVANT RESIDUES: (124)..(504)
US-09-796-487-4

Query Match 78.5%; Score 1582.5; DB 10; Length 381;
Best Local Similarity 79.6%; Pred. No. 8.9e-147;
Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

```
QY 7 PRGLPPRCRVLLVNLNPGCGKALOLFRRSHVQPLAEAEISFTLMTERRNHARELYRS 66
Db 6 PRGLPPRCRVLLVNLNPGCGKALOLFRRSHVQPLAEAEISFTLMTERRNHARELYCA 65
QY 67 EELGRMALVMSGDGLMHEVNGIMERPDMETAIORPLCSLPAGSGNALAASLNHYAGY 126
Db 66 EELGRMALVMSGDGLMHEVNGIMERPDMETAIORPLCSLPAGSGNALAASLNHYAGY 125
QY 127 EQVTNEDLLNCTLLLCRPVLSMNLISLHTASGLRFSVLSLWGLFADVDESXKRR 186
Db 126 EQVTNEDLLNCTLLLCRRRLSPMNLISLHTASGLRFSVLSLWGLFADVDESXKRR 185
QY 187 LGEHRTLTGFLRLAALRTYGRGLATLPVGRVGFRTSPAPVVOGPPVDALVPLEBVP 246
Db 186 LGEHRTLTGFLRLAALRTYGRGLATLPVGRVGFRTSPAPVVOGPPVDALVPLEBVP 244
QY 247 SHMOVDEDEFLVLTALSHLASEMFAAPMGRCAGVMHLYYRAGYSRALLRLFLAM 306
Db 245 SHMTVPPEQDFLVLTALSHLASEMFAAPMGRCAGVMHLYYRAGYSRALLRLFLAM 304
QY 307 EKGHRMEYECPLYLVVPVAFRLPEPKDGKGFVAVDGLMSEAVOGVHPNYFMVSGCV 366
Db 305 OKGRHMEIDCPYLHVVPVAFRLPEPKDGKGFVAVDGLMSEAVOGVHPNYFMVSGCV 364
QY 367 EPPSKRQOMPPEEP 383
Db 365 DAPSGRDSRRGPPPEEP 381
```

RESULT 10

```
US-09-796-487-5
: Sequence 5, Application US/09796487
: Patent No. US20020042358A1
: GENERAL INFORMATION:
: APPLICANT: Spiegel, Sarah
: TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
: FILE REFERENCE: 07320001aa (2033957-0001)
: CURRENT APPLICATION NUMBER: US/09/796,487
: CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: US 60/186,532
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: US 09/530,868
: PRIOR FILING DATE: 2000-05-05
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 5
: LENGTH: 373
: TYPE: PRT
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (1)..(373)
: OTHER INFORMATION: SEQ ID NO 5 is the peptide sequence of Mouse SPK1a in Fig. 2,
: OTHER INFORMATION: corresponding to amino acid residue 131 to 504 of SPK1a of Genb
: OTHER INFORMATION: quence Accession Number AAC61697.
: PUBLICATION INFORMATION:
: DATABASE ACCESSION NUMBER: AAC61697
: DATABASE ENTRY DATE: 1998-09-26
: RELEVANT RESIDUES: (132)..(504)
: US-09-796-487-5
```

Query Match 77.6%; Score 1564.5; DB 10; Length 373;
Best Local Similarity 79.4%; Pred. No. 5e-145;
Matches 297; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

QY 10 VLPRPCRVLYLNPGRGKAKALQTFRSHVPLLAETISFTLMTERRNARELYVEEL 69
DB 1 LIPRPRVLYLNPGRGKAKALQTFRSHVPLLAETISFTLMTERRNARELYVEEL 60
QY 70 GRWDALVMSGDGLMEHYVNGIMERPDMETALOKFLCSIPAGSGNALASLNHYAGYEOV 129
DB 61 GHWDALVMSGDGLMEHYVNGIMERPDMETALOKFLCSIPAGSGNALASLNHYAGYEOV 120
QY 130 TNEIDLNTCTLCRPLVSPMNLISLHTASGLRSFVSLAMGFIADVDLESCKYRRLGE 189
DB 121 TNEIDLNTCTLCRPLVSPMNLISLHTASGLRSFVSLAMGFIADVDLESCKYRRLGE 180
QY 190 MRETLGFLRLAALRTYKGLALPLPGRGVGFKTPRASPVVVVOGPDVAHLPLEEQVPSHW 249
DB 181 IRFVGTGFPRFLASIRTYQGLAYLPVGTASKRPAS-TLVQKGPVDTHTVPLEEVPSPHW 239
QY 250 QVVDDEPVLVALLHSHLASSEMPAPMGRCAGVNHLYVRAVSRAALLRLFLAMKG 309
DB 240 TVVEQDFLVLVLLHSHLASSEMPAPMGRCAGVNHLYVRAVSRAALLRLFLAMKG 299
QY 310 RHMEYCPVLYVYVAFLEPRDGRKGVAVDGLVSEAVVQGVPHNTFMVSGCEPDP 369
DB 300 KHMEDCPVLYVYVAFLEPRDGRKGVAVDGLVSEAVVQGVPHNTFMVSGCEPDP 359
QY 370 PSWKPOQMPPEEP 383
DB 360 SGDRSRGPPPEEP 373

RESULT 11

US-10-023-282-328
Sequence 328, Application US/10023282

Publication No. US20030092893A1

GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

FILE REFERENCE: P2007P1

CURRENT FILING DATE: 2001-12-20

EARLIER APPLICATION NUMBER: 09/205,258

EARLIER FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,884

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,964

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,882

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,899

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 328
LENGTH: 293
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (36)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-023-282-328

Query Match 73.8%; Score 1487; DB 9; Length 293;
Best Local Similarity 96.6%; Pred. No. 1,4e-137;
Matches 283; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 92 MERPDWETAIOKPLCSIPAGSGNALASLNHYAGYEQVNEIDLNTCTLCRPLVSPMN 151
DB 1 MERPDWETAIOKPLCSIPAGSGNALASLNHYAGYEQVNEIDLNTCTLCRPLVSPMN 60
QY 152 LLSLTASGLRSFVSLAMGFIADVDLESCKYRRLGEARFTLGTPLRLAALRTYGRRLA 211
DB 61 LLSLTASGLRSFVSLAMGFIADVDLESCKYRRLGEARFTLGTPLRLAALRTYGRRLA 120
QY 212 TLVGRGVGFKTPRASPVVVVOGPDVAHLPLEEQVPSHWQVVDDEPVLVALLHSHLASE 271

Db 121 YLPGRVSGSKTPASPVVYVQGPVDAHLVPLEEPVSHMTYVPPDEDFVTLALSHLASE 180
QY 272 MFAPMGCAGVHMLFYVRAGVSRAMLRLFLAMEKRHRHECPYLYVPPVAFRLRP 331
Db 181 MFAPMGCAGVHMLFYVRAGVSRAMLRLFLAMEKRHRHECPYLYVPPVAFRLRP 240
QY 332 KDGGVFAVDGLMVAEAVOGVHPNRYFMVSGCEPPPSMKPQOMPPEEPL 384
Db 241 KDGGVFAVDGLMVAEAVOGVHPNRYFMVSGCEPPPSMKPQOMPPEEPL 293

RESULT 12

US-09-796-487-9
; Sequence 9, Application US/09796487
; Patent No. US20020042358A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001a (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent version 3.1
; SEQ ID NO 9
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Putative kinase sequence obtained by assembling sequences from
; OTHER INFORMATION: several human ESTs (accession numbers D31133, AA232791, W63556, AA
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(204)
; OTHER INFORMATION: Corresponding to peptide sequence Putative Human in Figure 2.
US-09-796-487-9

Query Match 46.2%; Score 930.5; DB 10; Length 204;
Best Local Similarity 57.8%; Pred. No. 3.3e-83;
Matches 197; Conservative 4; Mismatches 3; Indels 137; Gaps 2;
QY 21 LNPGRGKALQFLRSHVQPLLAERAEISFTLTLTRRNHARELVASEELGRDALLVMSG 80
Db 1 LNPGRGKALQFLRSHVQPLLAERAEISFTLTLTRRNHARELVASEELGRDALLVMSG 60
QY 81 DGLMEVYVNGLMERPDMEFAIOKPLCSLPAGSGNALASLNHYAGYEQTNEDLLTCTL 140
Db 61 DGLMEVYVNGLMERPDMEFAIOKPLCSLPAGSGNALASLNHYAGYE----- 107
QY 141 LLCRPVLSPMNLLSLHTASGLRSFVLSLAWGFIADVLESDKYRRLGEMRFTLGTFLRL 200
Db 108 -----LFSVLSLAWGFIADVLESDKYRRLGEMRFTLGTFLRL 145
QY 201 AALRTYRGRLATLPYGRVGFETPPASPVVYVQGPVDAHLVPLEEQVPSHQVYPPDEDFVLY 260
Db 146 AALRTYRGRLATLPYGRVGSK----- 166
QY 261 LALLSHLASEMFAPMGCAGVHMLFYVRAGVSRAMLRLFLAMEKRHRHECPYLY 320
Db 167 ----- 166
QY 321 YVPVAFRLPEPKDGGVFAVDGLMVAEAVOGVHPNRYFMVSGCEPPPSMKP 361
Db 167 ---VAFRLPEPKDGGVFAVDGLMVAEAVOGVHPNRYFMVSGCEPPPSMKP 204

RESULT 13
US-09-817-676A-12
; Sequence 12, Application US/09817676A

; Patent No. US20020042101A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takatumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; TITLE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent version 2.0
; SEQ ID NO 12
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-817-676A-12

Query Match 43.4%; Score 875; DB 10; Length 617;
Best Local Similarity 38.9%; Pred. No. 4.4e-77;
Matches 187; Conservative 63; Mismatches 111; Indels 120; Gaps 4;
QY 10 YLPGRVYVLLNPGRGKALQFLRSHVQPLLAERAEISFTLTLTRRNHARELVASEEL 69
Db 141 LLPKRPRLIIVNPPGGGGLAMORCMHYVPMSISAGLSFNLTQTERGNHARELVQGLSL 200
QY 70 GRMDALVYVSGDGLMEHYVNGLMERPDMEFAIOKPLCSLPAGSGNALASLNHYAGYRQV 129
Db 201 SEMEGIVYVSGDGLIYEVYVNGLMERPDMEFAIOKPLCSLPAGSGNALASLNHYAGYRQV 260
QY 130 TNEDLTCTTLTLCRPVLSPMNLLSLHTASGLRSFVLSLAWGFIADVLESDKYRRLG 189
Db 261 VGVLDLLNCSLILKCGSHPLDLSVTLASGRCFSPLVAMGFLSDVDIHSERFALGS 320
QY 190 MFTLGTFLRLAALTYRGRLATLPV----- 215
Db 321 ARFTGAVYVGLASLTYRGRSLYPATTEPALPIGHSILPRAKSELVLAAPAPAAHTSP 380
QY 216 ----- 215
Db 381 LHRVSDLPPLPOPALVPSGSPEDLPDLSLNGGPELTGDMGAGDAPLSPDPLPSSP 440
QY 216 -----GRVGFETP--ASPVVYVQGPVDAHLVPLEEQVPSHQVYPPD 254
Db 441 NALKTAQSLPAEGPEPPASSGFLPPHSAPEASTGCVDDLPLPPLPQDQVTTIEG 500
QY 255 EDFVLYLALSHLASEMFAPMGCAGVHMLFYVRAGVSRAMLRLFLAMEKRHRHE 314
Db 501 E-FVLMGLITSLCADLMAAPHARFDGAVHLGVRSIGSIRALLRIFLAMEKHNFSL 559
QY 315 ECPYLYVPPVAFRLPEPKDGGVFAVDGLMVAEAVOGVHPNRYFMVSGCEPPPSMKP 374
Db 560 GCPHIGVAAARAFRLPEPLPRLTLVDELVEYGPLOAVHVGTLTLTG----PAGQKP 615
QY 375 Q 375
Db 616 Q 616

RESULT 14
US-09-817-676A-14
; Sequence 14, Application US/09817676A
; Patent No. US20020042101A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; APPLICANT: Kohama, Takatumi
; TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
; TITLE OF INVENTION: Expression and Methods of Use Thereof
; FILE REFERENCE: 00170/HG
; CURRENT APPLICATION NUMBER: US/09/817,676A
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/194,318

Mon Jul 14 09:16:40 2003

us-09-784-810a-2.rapb

Page 8

PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
US-09-817-676a-14

Query Match 42.5%; Score 856.5; DB 10; Length 618;
Best Local Similarity 38.6%; Pred. No. 2.9e-75;
Matches 189; Conservative 63; Mismatches 111; Indels 127; Gaps 6;

QY 10 VLPRPCRVLLNPNPGKAKALQLEFRSHVQPLLAELAEISFTLMLTERRNHARELYRSEEL 69
DB 140 LPRPRRLILLVNPFGGKGLAMQMKXHVLPMTISEAGLSFNLIQTERONHARELYVOGLSL 199
QY 70 GRMDALVWNSGGDLNHEVYNGIMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEQV 129
DB 200 SEMDGIYVSGDGLNHEVYNGIMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEQV 259
QY 130 TNEDLTNTCTLLCRPVLSPMNLISLFTASGLRSFVSLANGFTADVLDSDKTYRRLGE 189
DB 260 LGIDLNLNCSLLCCKGCGHPDLDLSVTLASGSCFSLVANGFTSDVDIOSERRALGS 319
QY 190 MREFTIGFLRLAALRTYRGLATLPV----- 215
DB 320 ARFTLGTVGLATLHTYRGLSYLPATVEPASPTPAHSILPRAKSELTLTPDPAPMAHSP 379
QY 216 ----- 215
DB 380 LHRVSQDLPLPLQPALASPGSPREPLTILSNGGPELAGDWGAGADAPLSPDLLSSPP 439
QY 216 --GRVGEKTPAS--PYVYQO-----GPVDAHLVPLEEQVPSHMOYVP 253
DB 440 GSPKALHSFVSEGAFTVTPSSGFLPPTPDARVAGASTCGPPDLPPLGTPLPPDM-VTL 498
QY 254 DEDFVLVLAISHLASEMFAAPMGRCAGVMILFYRACVSRAMLLRLFLAMEGRME 313
DB 499 EGDFTVLMALISPSHLAGADLVAAAPHARFDDGLVHLCWVRSISRAALLRLFLAMEGRSHFS 558
QY 314 YECPYLVYVVAFFRLPEKDGKGFVAVDGEIMVSEAVOGVHNTFMVYSGCVERPPPSWK 373
DB 559 LGCPLGYAARAARAFRLPELTPRGVLTVDGEQVEYGPLQAOHMPGIGTLTGTG---PPGC- 613
QY 374 PQQMPPEEP 383
DB 614 -----PGRREP 618

RESULT 15
US-09-970-516-4
; Sequence 4, Application US/09970516
; Patent No. US20020099029a1
; GENERAL INFORMATION:
; APPLICANT: No. US20020099029a1art1s AG
; TITLE OF INVENTION: Induction of blood vessel formation through administration of
; FILE REFERENCE: 4-31617
; CURRENT APPLICATION NUMBER: US/09/970,516
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-516-4

Query Match 42.5%; Score 856.5; DB 10; Length 618;
Best Local Similarity 38.6%; Pred. No. 2.9e-75;
Matches 189; Conservative 63; Mismatches 111; Indels 127; Gaps 6;

QY 10 VLPRPCRVLLNPNPGKAKALQLEFRSHVQPLLAELAEISFTLMLTERRNHARELYRSEEL 69
DB 140 LPRPRRLILLVNPFGGKGLAMQMKXHVLPMTISEAGLSFNLIQTERONHARELYVOGLSL 199
QY 70 GRMDALVWNSGGDLNHEVYNGIMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEQV 129
DB 200 SEMDGIYVSGDGLNHEVYNGIMERPDMETAIQKPLCSLPAGSGNALAASLNHYAGYEQV 259
QY 130 TNEDLTNTCTLLCRPVLSPMNLISLFTASGLRSFVSLANGFTADVLDSDKTYRRLGE 189
DB 260 LGIDLNLNCSLLCCKGCGHPDLDLSVTLASGSCFSLVANGFTSDVDIOSERRALGS 319
QY 190 MREFTIGFLRLAALRTYRGLATLPV----- 215
DB 320 ARFTLGTVGLATLHTYRGLSYLPATVEPASPTPAHSILPRAKSELTLTPDPAPMAHSP 379
QY 216 ----- 215
DB 380 LHRVSQDLPLPLQPALASPGSPREPLTILSNGGPELAGDWGAGADAPLSPDLLSSPP 439
QY 216 --GRVGEKTPAS--PYVYQO-----GPVDAHLVPLEEQVPSHMOYVP 253
DB 440 GSPKALHSFVSEGAFTVTPSSGFLPPTPDARVAGASTCGPPDLPPLGTPLPPDM-VTL 498
QY 254 DEDFVLVLAISHLASEMFAAPMGRCAGVMILFYRACVSRAMLLRLFLAMEGRME 313
DB 499 EGDFTVLMALISPSHLAGADLVAAAPHARFDDGLVHLCWVRSISRAALLRLFLAMEGRSHFS 558
QY 314 YECPYLVYVVAFFRLPEKDGKGFVAVDGEIMVSEAVOGVHNTFMVYSGCVERPPPSWK 373
DB 559 LGCPLGYAARAARAFRLPELTPRGVLTVDGEQVEYGPLQAOHMPGIGTLTGTG---PPGC- 613
QY 374 PQQMPPEEP 383
DB 614 -----PGRREP 618

Search completed: July 12, 2003, 18:12:31
Job time : 137 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 18:02:02 ; Search time 28 Seconds

(without alignments)
1318.415 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGSGRGVLPKRCRYLV.....CVEPPSGMKPQOMPPEEPL 384

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	438	21.7	1240	2	T05162
2	372.5	18.5	473	2	T19707
3	369	18.3	458	2	T38776
4	355.5	17.6	687	2	S51398
5	337	16.7	624	2	S67059
6	233.5	11.6	549	2	T33517
7	141.5	7.0	306	2	AH1769
8	133	6.6	302	2	F86849
9	132.5	6.6	309	2	H65995
10	130.5	6.5	303	2	F69795
11	129.5	6.4	306	2	AE1394
12	125	6.2	295	2	A83894
13	123.5	6.1	311	2	F83871
14	120.5	6.0	310	2	A61665
15	120.5	6.0	364	2	F84898
16	118	5.9	309	2	AH1528
17	115	5.7	315	2	AB2166
18	114	5.7	342	2	D86677
19	113	5.6	309	2	AF1171
20	111	5.5	294	2	G95120
21	111	5.5	311	2	C97990
22	111	5.5	321	2	C70596
23	108.5	5.4	310	2	A11293
24	108	5.4	315	2	A89978
25	106.5	5.3	433	2	S75948
26	106	5.3	309	2	H70861
27	104	5.2	304	2	F72386
28	103.5	5.1	295	2	D83734
29	100.5	5.0	311	2	D75405

30	100	5.0	1273	2	E72611	probable ATP-depen
31	97.5	4.8	305	2	G89844	hypothetical prote
32	97	4.8	449	2	C83634	hypothetical prote
33	96.5	4.8	475	2	H84567	probable diacylgly
34	94.5	4.7	254	2	G75333	3-oxoacyl-acyl car
35	94.5	4.7	260	2	G87400	conserved hypothet
36	94	4.7	1247	2	A33812	interphotoreceptor
37	93	4.6	297	2	F69595	multidrug resistanc
38	92.5	4.6	732	2	T16422	hypothetical prote
39	91.5	4.5	317	2	T00550	probable phospholi
40	91.5	4.5	881	2	F83530	hypothetical prote
41	90.5	4.5	304	2	A86842	conserved hypothet
42	90.5	4.5	392	1	A70406	transporter (major
43	90.5	4.5	1346	2	T17412	polyketide synthas
44	90.5	4.5	1694	2	A83512	hypothetical prote
45	90	4.5	260	2	AF3374	hypothetical prote

ALIGNMENTS

RESULT 1

T05162
hypothetical protein F185.160 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

C:Accession: T05162

R:Byron, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.;

submitted to the Protein Sequence Database, August 1998

A:Reference number: 215400

A:Accession: T05162

A:Molecule type: DNA

A:Residues: 1-1240 <BEV>

A:Cross-references: EMBL:AL022603

A:Experimental source: cultivar Columbia; BAC clone F185

C:Genetics:

A:Map position: 4

A:Introns: 44/3; 117/3; 178/3; 206/1; 237/2; 375/1; 416/1; 449/3; 474/1; 503/1; 529/1

A:Note: F185.160

Query Match 21.7%; Score 438; DB 2; Length 1240;

Best Local Similarity 31.5%; Pred. No. 6.1e-30;

Matches 119; Conservative 65; Mismatches 132; Indels 62; Gaps 13;

QY	11	IPRCRVYVLPNPRGKGRKALDERSHOPPLAEAFISFTLMTERRRHARLVASEELG	70
DB	374	IGRPKRLVFNVPFGKRSAREIFVEKVPLEFDADVOLEIQETYYQLHAKEFVKSMDVS	433
QY	71	RMDALVVASGDGLMHEVNVNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGEQVT	130
DB	434	KYDGIVCVSGDILVEVNVNGLMERADWNAKLPIGMVPAAGNGMISLDTVGLRCCA	493
QY	131	NEDLLTNCTLLCRPVLSPMNLSLHTASGLSFVSLSLWGFADVDLSDKYRLGEM	190
DB	494	N-----SATISITIRHKRSVDVATL-AOGNTKFFSVLMLAWGLADIIDIESKFRMGSA	547
QY	191	FTLTGTLF-----RLAALFTYGRLATLPV-GRVQKFPASPVVQV-	230
DB	548	RIDFVYCLVDKFDNVCIAVVKLLALORITCLRRYNGRIILFPAPEFGYGGPASCLSLQOE	607
QY	231	-----OGPVDAHLVPLEEGVPSHWQVPPDDFVTLALLHS--HLASEMFAAPMG	278
DB	608	PHVSKVEGYGQP-ETKREDFE-----WREKGP---FVTITLNVNMGSENTLTAPAA	657
QY	279	RCACGVMLFFYVRAGVSRAMLRLFLAMEKGRHMEYEGPYLVV-----PVVAFRLP-	331
DB	658	KFSDGYLDLIVLK-NCPLVILSLMRQSSGTHV--ESPYYIYIKLVEKVAFLERGA	714
QY	332	-----KDGKGVAVDGEELM	345
DB	715	LVDEPDKRGIIIDSDGEVL	732

RESULT 2

hypothetical protein C34C6.5 - *Caenorhabditis elegans*
T19707

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T19707

R:Percy, C.
submitted to the EMBL Data Library, October 1995

A:Reference number: 219167

A/Accession: T19707

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-473 <MIL>

A/Cross-references: EMBL:Z66494; PIDN:CAA91259.1; GSPDB:GNO0020; CESP:C34C6.5

A/Experimental source: clone C34C6

C/Genetics:

A:Gene: CESP:C34C6.5

A:Map position: 2

A/Introns: 82/1; 126/1; 158/3; 276/1; 311/3; 427/3

Query Match 18.5%; Score 372.5; DB 2; Length 473;

Best Local Similarity 28.2%; Pred. No. 9.5e-25;

Matches 109; Conservative 87; Mismatches 139; Indels 51; Gaps 14;

```

QY 12 PRP---CR--VLVLNPRGKGKALQFRSHVQPLAEAFISFTIMTERNHARE-LVR 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 PPQEQGRGNLIVINPNSGSKLEFFANTVGGKLDKSLRYEVVTTGPNHARVLMT 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 66 SEELGRDALVYMSGDLMEHVVNGLMERPQ-WETAIOKPLCSIPAGSGNALAAS-LNH 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 KADGKENGVLILSGDGLVEALNGILCRDPAFRIFPLDIGIVPSGSGGLCSVLSKY 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 124 AGYQVTVNEDLTNCTLLCRPVLSPMNLISLTASGLRSP-SVLSLAPFIADVDESD 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 G--TKMKEKSMERALIATSPFAKASVALISVKTNOASYSLISGGLMADIDISE 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 183 KYRR-LGEMRPTLCTFLRLAALTYRGRLATLPVGRVGFKTPASPVV----VQGG 236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 KMRSLGHHRFTVGFIRSCNLIRYKGRILYRYPKPGFH-PSANVSVEYKTTQGRID 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 237 HLYVL-----EEQVPSHMQ-----VYPPDEDFVLVALLASH 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 SKVATNGSVSDSEETMETKQNTLPDSDETLAVSSDLEETVVIDENVTNITAVLSH 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 268 LASEMPAPMGRCAGVYMLFYV--RAGVSRAMLRLFLAMEGRHMEYECPLYVYVV 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 375 IADGPRAPSAKLEDNRILHSYLIMKDIGTRVNIATKYLALIEHETHL--DLPEVK 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 AFRLPEKDKGVAVNDELAVSEAVQ 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 SMKLEVISGSHVVLDEVDYTKTIE 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3

hypothetical protein SPAC4A8.07c - fission yeast (*Schizosaccharomyces pombe*)
T38776

C:Species: *Schizosaccharomyces pombe*

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C/Accession: T38776

R:Skellton, J.; Churche, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997

A:Reference number: 221751

A/Accession: T38776

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-458 <SKE>

A/Cross-references: EMBL:Z99762; PIDN:CA11477.1; GSPDB:GNO0066; SPDB:SPAC4A8.07c

A/Experimental source: strain 972h-; cosmid c4A8

C/Genetics:

A:Gene: SPDB:SPAC4A8.07c

A:Map position: 1

A/Introns: 39/1; 101/1

Query Match

18.3%; Score 369; DB 2; Length 458;

Best Local Similarity 28.6%; Pred. No. 1.8e-24;

Matches 106; Conservative 67; Mismatches 142; Indels 56; Gaps 10;

```

QY 11 LPRPCRYVLLNPRGKGKALQFRSHVQPLAEAFISFTIMTERNHARELVSEELG 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 IKRSRRFIYVINHGGKAKHWESEAPFESSHISCEVVLTRKRDAKSIATKMDVG 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 71 RMDALVMSGDLMEHVVNGLMERPQWETAIOKPLCSIPAGSGNALAASLNHAGVEYQT 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 SYDGLISVGADGFLHEVINGLGERDYLEAFKLPVCMIPGSGNAFS-----YATGQL- 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 131 NEDLLTNCITLLCRPVLSPMNLISLTASGLRSPSVLSLAPFIADVDESDKYRRLGEM 190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 216 -KPAALALELTKRP--TSPDLMTFE-QGKKAYSLFLANYGIIADCDIGETEMWRMGEM 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 191 RFTLGTFLRLAALTYRGRLATLPVGRVGFKTPASVYVQGGVDA-----H 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 RAYLGFFRL-----FQKPMKCSIEMDVYSSDRTIEIKHYEKSNN 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 238 LVPLEE-----QVPSHMQVPPDEDFVLVALLSHLASEMPAPMGRCAGVML 287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 LAPMSESSDSDKTVSTSPESHLLTFEINDISIFCAGILPIADAMFPAASNDSDIIV 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 288 FYRAGVSRAMLRLFLAMEGRHMEYECPLYVYVVAERLEP-KDGR-GVPAVDGELM 345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 VIYYSQFRRKSLSMFTQDNGF--YSHLNYVYRSRFTPVMTGRHYRFPALDGESY 430
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 346 VSEAVQGGVHP 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 431 PLEPFECRVAP 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4

hypothetical protein YLR260w - yeast (*Saccharomyces cerevisiae*)
S51398

N/Alternate names: hypothetical protein L8479.7

C/Species: *Saccharomyces cerevisiae*

C/Date: 05-May-1995 #sequence_revision 12-May-1995 #text_change 19-Apr-2002

C/Accession: S51398

R/Miller, N.
submitted to the EMBL Data Library, November 1994

A/Description: The sequence of S. cerevisiae cosmid 8479.

A/Reference number: S51395

A/Accession: S51398

A/Molecule type: DNA

A/Residues: 1-687 <MIL>

A/Cross-references: EMBL:U17244; NID:g5771171; PIDN:AA67377.1; PID:g577178; GSPDB:GNO

C/Genetics:

A:Gene: SGD:LCB5; MIPS:YLR260w

A/Cross-references: SGD:S0004250

A/Map position: 12R

Query Match 17.6%; Score 355.5; DB 2; Length 687;

Best Local Similarity 24.9%; Pred. No. 4.9e-23;

Matches 102; Conservative 76; Mismatches 150; Indels 81; Gaps 11;

```

QY 17 VLVLNPRGKGKALQFRSHVQPLAEAFISFTIMTERNHARELVSEELGMDALV 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 IFVILNPFGRGKAKAKFMFKARPKILLASCSIEVYTKPGALIELAREMDIDKDTIA 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 77 VMGDDGLMEHVVNGLMERPQWETAIOK-PLCSIPAGSGNALAASLNHAGVEYQTNEDLL 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 CASGDGIPHVINGLYORPDHVAFAFNIAITIEIPCGSGNAMSVC-HM-----TNPN-- 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 TNCITLLCRPVLSPMNLISLTASGLR--SPVLSLAPFIADVDESDKYRRLGEMR 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 SYSTLCIKIKIEFRIDLMCCSOPSYAREHPKLSFLSQTGLLAETDINTIEFIWMGPARE 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 193 TLGTFRLAALRTY-----NGRLATLPV----- 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 ELGYAFANIOKKKTYPCETIYYKAAKSNELKNHYLEHKNKGSLPEQHTITNKNKNDODNY 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Oy      216 ---GRGFKTPASPVVVOGGPVDAHV-----PLEEQVPSHMQV 251
Db      502 NYENYEYETEDEDDEDADDEDSDHLSISDLADSSAQIKEEDPKIRYPLDEGIPSDMER 561
Oy      252 VPDEDPVVLVALLHS---HLASEMFAAPMGRCAAGVHHLLYVAGVSRAMLLNLFNAME 307
Db      562 L-DPNISNNIGIIFYTGKMPVADRPKFPALPSGTDMDVITDAPRSLTRMADPILGLD 620
Oy      308 KGRHMEECPCYLVVYVVAFAFLPEPKDGKGFVAVDELAVYSAVGOVHP 356
Db      621 KGSHVLD--PEVYLHSKILAYKIIRPKLGNGLPSVDSGEKFLPELPLOYEIMP 667

```

RESULT 5
S67059
hypothetical protein YOR171c - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein 03615
C:Species: *Saccharomyces cerevisiae*
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S67059
R:Bordone, R.; Camasses, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarassov, I.A.; Winc
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67032
A:Accession: S67059
A:Molecule type: DNA
A:Residues: 1-624 <BOR>
A:Cross-references: EMBL:Z75078; NID:g1420415; PID:g1420417; GSPDB:GN00015; MIPS:YOR171d
A:Experimental source: strain S286C
C:Genetics:
A:Gene: SGD:ICB4; MIPS:YOR171c
A:Cross-references: SGD:S0005697
A:Map position: 15R

	Query Match	16.7%	Score 337;	DB 2;	Length 624;
	Best Local Similarity	25.3%;	Pred. No. 1.8e-21;		
	Matches	100;	Conservative	82;	Mismatches 139; Indels 74; Gaps 14.
QY	17	VLVLLNPRGGKGAQLFRSHVOPLLAAEASFTLMLTERHNRHARELYRSEELGRMDLV	76		
Db	229	ILVILNPHGGKGTAKNLEITKARPIIVSGSGKIEIAYIKRYARHAIIDAKIDLDISDYDTA	288		
QY	77	VMSGDGLMHEVYVNGIMERBDWETAIQK-PICLSIPAGSGNALAASLNHTAGYEQVYNEDLL	135		
Db	289	CASGGGLPEYEYVINGLYRRPDRVDANKLAIVQLPGCGSSNAMSISC-HW-----TNNP--	339		
QY	136	TNCTLLLCRPLSPNNLLSLTASGLNSF-----SVLSLWGFIAVDLDESCKYRRGEMRF	192		
Db	340	SYALCLIVKSIETRIDLMCCSQSPYMNEMPRLSQFYGVYAESDINTFEIRFMGVPYRF	399		
QY	193	TLG-----TFRLAA-----LRTY-----RGRATLTPVGRVGRKTP--	223		
Db	400	NLGAFANIIOGKKYPCFEVFRYKAAKSKKELVHLENDKNKGCTLFFP-----NPSF	452		
QY	224	-ASPVVVOQGQVDAHL-----VLEQVQPSHMQVVDDEDFVLVLLAH	265		
Db	453	NSSDDLSTKNNINNSTKDELSPNLENDENFKLYKYTEPVRPDWCKM-DSELDLTMLTFY	511		
QY	266	S-----HLASEKFAAPMGRCAAGVNHLYTVRAGVSRAMLLRLFLAMEKGRHNEYECPLYLV	321		
Db	512	TGKMPYIKDRKFFPAALPADGTIDLVTYDARIPTVTRPTPIILSDKSHYLE--PEYIH	569		
QY	322	VPVVAFLRLEPKDQGVFAVDSELWASLVAQGVHP	356		
Db	570	SKILLAYKTIIPVSGESLVEGDKRPFLEPLQVQIIMP	604		

RESULT 6
T33517
hypothetical protein T10B11.2 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33517
C:MLinx, P.J. Kemp, K.

submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid T10B11
A:Reference number: Z21363

A:Accession: J1351/ translated from GB/EMBL/DDBJ
 A:Status: preliminary;
 A:Molecule type: DNA
 A:Residues: 1-549 <MTN>
 A:Cross-references: EMBL:AF098993; PIDN:MAC6746.1; GSPDB:GN00019; CESP:T10B11.2
 A:Experimental source: strain Bristol N2; clone T10B11
 C:Genetics:
 A:Gene: CESP:T10B11.2
 A:Map position: 1
 A:Introns: 26/1; 76/2; 109/3; 159/1; 187/3; 229/3; 353/1; 398/2; 434/1; 467/3

```

Query Match          11.6%: Score 233.5; DB 2; Length 549;
Best Local Similarity 24.1%: Pred. No. 1.9e-12;
Matches 100; Conservative 72; Mismatches 150; Indels 93; Gaps 19;

OY      8  RGVLPRCCRVLVLLNPPGKGKGLQQLFRSHVQPLL-AGEAISTFLMTERRNHARELVRS 66
DB      158 RRVRNRKPKNIIFINPPGNGKAKOKIFKDVNDVAFFWLTPGLRKRVVLTERRANHARQIVE 217
OY      67  EELGRW---DALVMSDDGLMHEVYNQIMERPMETA-IGKPLCSL-----PAG 111
DB      218 MPEWMSAIDSLVSVGGDGLFNEILSGLLTQTDAGRINDNPSHLVTPHIRRGITAG 277
OY      112 SGNALASLNNFYASGEVQTNED-----LITNCTLLCRPYLSPMNLSTLTASGLRS 163
DB      278 SANISVSYVHE-----TNDHATSAVHAIAIGSCNNVDC-----TVHQHKILR 320
OY      164 FSVSLAMGFIADVDDLESKYRRLGEMKRFYLTGTFLLALALRT-----YGRLATLPEVR 217
DB      321 ISANAISYGMIGDVLIRSEERYRCGLPIRY-----QMSALRTTIRHPIYR-----GM 366
OY      218 VGFKTPASPVYVQGGPVDANLVPLREQV-----SHNQVVRDDEDFVLYALALH 265
DB      367 VQFSLSIKENV---NPKD-QLPRLCEPCRYCMKPGCKDKDYKHNA-----EFTHVICVI 418
OY      266 SHLA-SEMFAPMGRCAGAYMHLFYV-RAGVSRHMLRLFLAMEKGRHMEYECOPYLVV 322
DB      419 PTVYPTFPYGLAPFTGIDGFTLDLALVPRISRFHNMQFMKRVAYMGKQLYELDPSLNCY 478
OY      323 PVAFARLEP---KQSKGFVANVDGELM---VSEAVQGVNPR---NYFMVNGGCEP 368
DB      479 RVTWMSYQPRDADQDQDPGVWNLIDGELIDQRPDEPLHFLHQLDLSFEGRDAAWKP 533

```

RESULT 7
 AH1769
 conserved hypothetical protein lin2702 [imported] - *Listeria innocua* (strain C1p1126)
 C:Species: *Listeria innocua*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AH1769
 R:Glaser, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgst, O.; Entlian, K.D.; Eshti,
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madeno, E.; Maitournam, A.;
 Ok, C.; Schlueter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AH1769
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-306 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97928.1; PID:q16415238; GSPDB:GN00178
 A:Experimental source: strain C1p11262
 C:Genetics:
 A:Gene: lin2702

 Query Match 7.0%; Score 141.5; DB 2; Length 306;
 Best Local Similarity 20.6%; Pred. No. 0.00011;
 Matches 73; Conservative 64; Mismatches 128; Indels 89; Caps 15;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 18:06:42 ; Search time 19 Seconds

(without alignments)
594.653 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGRGVLPKRCVRLV.....CPEPPSGMKPQMPPEPL 384

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	5.2	324	US-09-134-001C-4774	Sequence 4774, Ap
2	102.5	5.1	1065	US-09-412-545-2	Sequence 2, Appl1
3	96	4.8	712	US-09-708-426-9	Sequence 9, Appl1
4	92	4.6	3782	US-09-105-537-4	Sequence 4, Appl1
5	91	4.5	2584	US-08-936-135-4	Sequence 37, Appl1
6	90.5	4.5	1346	US-09-105-537-37	Sequence 6, Appl1
7	90.5	4.5	11877	US-09-105-537-6	Sequence 2, Appl1
8	90	4.5	2588	US-08-936-135-2	Sequence 2, Appl1
9	88	4.4	6095	US-09-144-085-2	Sequence 5553, Ap
10	87.5	4.3	319	US-09-134-001C-5553	Sequence 4, Appl1
11	87.5	4.3	1346	US-09-320-878-4	Sequence 2, Appl1
12	85.5	4.2	1421	US-09-335-409-2	Sequence 2, Appl1
13	85.5	4.2	1421	US-09-568-102-2	Sequence 2, Appl1
14	85.5	4.2	1421	US-09-568-969-2	Sequence 2, Appl1
15	85.5	4.2	1421	US-09-568-480-2	Sequence 2, Appl1
16	85.5	4.2	1421	US-09-568-486-2	Sequence 2, Appl1
17	85.5	4.2	1421	US-09-568-472-2	Sequence 2, Appl1
18	85.5	4.2	1421	US-09-567-899-2	Sequence 2, Appl1
19	84	4.2	3011	US-08-453-552-2	Sequence 2, Appl1
20	84	4.2	3011	US-08-710-637-2	Sequence 2, Appl1
21	84	4.2	3011	PCT-US93-00907-2	Sequence 2, Appl1
22	83	4.1	1011	US-08-850-328-5	Sequence 5, Appl1
23	83	4.1	3011	US-08-188-281B-1	Sequence 1, Appl1
24	83	4.1	3011	US-08-433-352-1	Sequence 1, Appl1
25	83	4.1	3011	US-08-710-637-1	Sequence 1, Appl1
26	83	4.1	3011	US-08-811-566-20	Sequence 20, Appl1
27	83	4.1	3011	US-09-014-416-1	Sequence 1, Appl1

28	83	4.1	3011	US-09-014-416-5	Sequence 5, Appl1
29	83	4.1	3011	US-08-850-328-1	Sequence 1, Appl1
30	83	4.1	3011	US-09-034-756-20	Sequence 20, Appl1
31	83	4.1	3011	PCT-US93-00907-1	Sequence 1, Appl1
32	83	4.1	3011	PCT-US94-07280-1	Sequence 1, Appl1
33	83	4.1	3011	PCT-US95-01087-1	Sequence 2, Appl1
34	83	4.1	3012	US-08-811-566-2	Sequence 2, Appl1
35	83	4.1	3012	US-09-034-756-2	Sequence 5, Appl1
36	82	4.1	390	US-09-192-983-5	Sequence 11, Appl1
37	82	4.1	435	US-08-031-538-11	Sequence 15, Appl1
38	81.5	4.0	2787	US-09-245-041-15	Sequence 31, Appl1
39	81	4.0	403	US-08-665-259-31	Sequence 31, Appl1
40	81	4.0	403	US-08-762-500-31	Sequence 17, Appl1
41	81	4.0	515	US-09-413-814-22	Sequence 2, Appl1
42	81	4.0	2556	US-08-185-432-17	Sequence 2, Appl1
43	81	4.0	2556	US-08-899-232-2	Sequence 2, Appl1
44	80	4.0	403	US-08-665-259-30	Sequence 30, Appl1
45	80	4.0	403	US-08-762-500-30	Sequence 30, Appl1

ALIGNMENTS

RESULT 1
US-09-134-001C-4774

; Sequence 4774, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

; FILE REFERENCE: GTC-007

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4774

; LENGTH: 324

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

; US-09-134-001C-4774

Query Match 5.2%; Score 105; DB 4; Length 324;
Best Local Similarity 19.1%; Pred. No. 0.0026;

Matches 70; Conservative 61; Mismatches 125; Indels 110; Gaps 18;

QY	16	RVLVLLNRGRGKALOLFRRSHVOPLLAEISFTLMTERNNHARELVRSEELG-----	70
DB	12	RRRIYNTSG-----ELFKR-----VLPDA-----LIRLEKAGETSAVATEKIGDQATFE	58
QY	71	-----KMDALVWMSGDGLMHEVYNGMERPDWETAOKPLCSLPAGSGNALASLNHY	123
DB	59	AERALESEYDILLIAGAGGTINEVVGNAEQPN-----RPKRGVIMGTVDQFALH--	111
QY	124	AGYEVTNEDLLTNTLLCRPVLSPPN-LLSLHTASLSRFSVLSLAMEFIADVDESD	182
DB	112	-----LPSDIMGAIIVIIDGHTTK-----VDIGKMN	138
QY	183	KYRRIGEMRFTLGPELRLAALRTYRGLATLPVGVGFTTPASPVVOGSPVDALHVLLE	242
DB	139	RY-----FINLAA-----GKLT-----OVSTETP-SKLKSTVGP-AYINGF	175
QY	243	EOVPS-----HMQVDEDEFLVALIALSHLASEMFAAPMRCACAGYMLFYVRA	292
DB	176	EMLPQMKAVDVRIEVDNIFQGEALLFILGLTNSAGEXL--VPPAKLDDGYFTLLIYEK	234
QY	293	GVSRRMLRLFLAMEKGRHMEYECPLYVVPVAFRLPEPKDGKGVFAVDGLMSEAVOG	352
DB	235	-ANLAEIHLHIMTLASRGHTKH--PKVIYAKASINISS-----FTDMQLNDGGEYG	284

QY 353 QVHPNY 358.
Db 285 KLIPANF 290

RESULT 2

US-09-412-545-2
Sequence 2, Application US/09412545
Patent No. 6255095
GENERAL INFORMATION:
APPLICANT: Prescott, Stephen M.
APPLICANT: Ding, Li
APPLICANT: Trier, Elle
TITLE OF INVENTION: HUMAN DIACYLGLYCEROL KINASE IOTA
FILE REFERENCE: 1321.2.25
CURRENT APPLICATION NUMBER: US/09/412.545
CURRENT FILING DATE: 1999-10-05
EARLIER APPLICATION NUMBER: 60/103.079
EARLIER FILING DATE: 1998-10-05
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1065
TYPE: PRT
ORGANISM: Homo sapiens
US-09-412-545-2

Query Match

Best Local Similarity 5.1%; Score 102.5; DB 4; Length 1065;
Matches 57; Conservative 30; Mismatches 61; Indels 81; Gaps 12;

QY 3 PAGPRGVLPKPCRVLVILNPRGSGK--KALQLFRSHVOPPLAEATISFTMLTERNNH 59
Db 368 PISP--LHKP--LVFVNPCKSGNGTKYLOKFMKYNL-----403
QY 60 ARELVRSEELGRMDAL-----VNSGDGIMHEVYNGIMERPDMETAIQKPLCSL 108
Db 404 -RQVFDLSQEGPKKALDYLRKVPMLRIACGGDGTVMILSIDE--LQSPQPVGV 459
QY 109 PAGSGNLAASLNHYAGYEQVNTNEDLTNCTLLCR-----PVLSP 149
Db 460 PLGTGNDLARTLNMGGST--TDEPV--SKILQVEDGTVOVOLDKMNILHVERNPDLPP 512
QY 150 MNL-----LSLHTASGLRSFSVLSLWGFADVDLESDKRYRLGEMRF 192
Db 513 EELEDGVCKLPLNFYNNY--FSL-----GFDHHTLTFHESREANPEKF 554

RESULT 3

US-09-708-426-9
Sequence 9, Application US/09708426
Patent No. 6444429
GENERAL INFORMATION:
APPLICANT: HAN, YE-SUN
APPLICANT: YU, YEON-GYU
APPLICANT: LIM, JAE-HWAN
TITLE OF INVENTION: GENE CODING FOR DNA LIGASE OF HYPERTHERMOPHILIC BACTERIA AQUIFEX
FILE REFERENCE: 199579USO
CURRENT APPLICATION NUMBER: US/09/708.426
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: KR99-49591
PRIOR FILING DATE: 1999-11-10
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Version 3.0
SEQ ID NO 9
LENGTH: 712
TYPE: PRT
ORGANISM: Rhodothermus marinus
US-09-708-426-9

Query Match 4.8%; Score 96; DB 4; Length 712;
Best Local Similarity 22.1%; Pred. No. 0.093;

Matches 64; Conservative 37; Mismatches 108; Indels 80; Gaps 11;

QY 59 HARLIVSEELGRMDALVWMSGDGLMHEVYNGLM-----ERPDMETAIQKPLCS-----107
Db 401 HNEIDYVRSRDRIQDLVIVIRAGDVIPQVYKPVVEARTGNERP--WRPEKCPSCGSQLVR 459
QY 108 LPAGSGNLAAS-----LNHYAGYEQVNTNEDLTNCTLLC-----RPV-----LS 148
Db 460 LPGEADYYCVASDCPPADPFLHEHFAGRDAMDITGMSQVAKOLAEGLYRPLSDLYRLK 519
QY 149 PMNLSLHTASGLRSFSVLSLWGFADVDLESDKRYRLGEMRFITGLRLAALRTYRG 208
Db 520 LEDLTKLEGFAETRAIRMLRA-----IEASKQPLSRILFGLG-----IRHYGK 563
QY 209 RLATLPYGRGFKTPASPVYVQO-----GPVNAHLP-----240
Db 564 TTAELVQRFASIDELAAATIDELAALEGVPIAESIAMFVEDNRRRIIEELKELGVN 623
QY 241 ---LEEQVPSHMQVPDEDFYLVALLH--SHLASEMFAAPMGRCAAGV 284
Db 624 TORLEAPRAESPVKGTFTVLTGALPHLTRKEAEELIKRAGGVAASSV 672

RESULT 4

US-09-105-537-4
Sequence 4, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 3782
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-4

Query Match

Best Local Similarity 4.6%; Score 92; DB 4; Length 3782;
Matches 99; Conservative 47; Mismatches 196; Indels 84; Gaps 17;

QY 3 PAGPRGVLPKPCRVLVILNPRGSGKALQLFRSHVOPPLAEATISFTMLTERNNHAR 61
Db 132 PEGGQRAVVEGIIHETLEGAPDPSASYAFELGCFVRPAVTAANAALVGPADRADFA 191
QY 62 ELVR-----SEELGRMDAL--VNSGDGIMHEVYNGIMERPDMETAIQKPLCSLPAG--SG 113
Db 192 DLLERLRLPSLSLAPQSLRTVRAADGALAEITLALDSDSDSGALLSALGVTAAYVLTG 251
QY 114 NALAASLNHYAGYQVYNTNEDLTNCTL-----LLCRPLSPMLNLSLHTASGLR 162
Db 252 NAVIALIAHPQWRELCDDRGGLAAVAEETRLRDPVQDARVVRGETELAGRLRPGAH 311
QY 163 SFSVLSLWGFADVDLESDKY--RRLGEMRFTL-----GTFRLT--AALRTYR 207
Db 312 -VVVLTAAATGRDEPVFTDPERFDLARPDAANHLALHAGYGVPAASLYRQAEVALRTLA 370
QY 208 GLATL-----PYGRVGFKTPASPVYVQGPVDAHLVPLEEQVPSHMQVVPDE 255
Db 371 GRFPGLRQAGDVLRPRAPVGRGPLSVYSSSMKVVLTSPAHHTHYGLVPLAMALLAAG 430
QY 256 DFVLYL---ALHSHLASEMFAAPMGRCAAGVNAHLYFR--AGYSRAALLFLAMEGRH 311
Db 431 HEYKVASQPALTDITITSGLAAPVVG--TDHLIHEYVRMAAGEBP-----NH 476
QY 312 MEYECPLVYVVPVAF---RLPEKDGKGVAVDGEIWLVSBAVOGVYPMYFMVSGCVER 368

Db 477 -----PALAFDAREPELDMCHALGIEAIL-----APYHLLANNDSMDLDYDF 521

Qy 369 PPSMKP 374

Db 522 ARSWOP 527

RESULT 5

US-08-936-135-4

; Sequence 4, Application US/08936135

; Patent No. 6054293

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: He, Zhigang

; TITLE OF INVENTION: Semaphorin Receptors

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/936,135

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC97-288-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 343-4341

; TELEFAX: (650) 343-4342

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2584 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-936-135-4

Query Match 4.5%; Score 91; DB 3; Length 2584;

Best Local Similarity 22.0%; Pred. No. 2.6; Mismatches 115; Indels 78; Gaps 13;

Matches 67; Conservative 45;

Db 41 LLAELISFTLMTERRNHARE--LYRSEELGRMDALVMS--GDGLMHEVNGMLMRP 95

Qy 2049 LYSEHHSVALGLYTHREAGCAVALYSLHISTYRGINLSPKGLITFASPEL----- 2104

Db 96 DWETAIQKPLCSLPAGSGLNALASLNHYAGE-----QVTNEDLLTNTCLLCRPVLS 149

Qy 2105 NLEVALTRPMETVA-----LVALGLYHISGLNGLYASPHISIRPLYSGLGYARGVALL 2159

Db 150 MNLSLHHTASGLRSTSVSLAMGFLADVLDSDKYRLG-----EMRTTLGTF-LRLAAL 203

Qy 2160 LEHISLVSSERLLEYSL-----ETRYRGLNVALILPEHNGSLGYGLILEGL 2204

Db 204 RTYGRRLAT-----LPVGRVGFETPPASPVVVOGQPYDA-----HLVPLEQO----- 244

Qy 2205 YLYSGLYASNEGLYGLYILALAVALASPAPIESRILEASASNSHISILEPRGLNG 2264

Db 245 --VPSHMVQVDEDFVLLALLSHSLASEMFAAPRGCAAGMHLFYVRAGYSRAMLR 302

Qy 2265 LASPCY-----SALALYSPTRTGHRASPLEASP-----LVSLYSASMTWHRGLIEL 2308

Qy 303 FLAME 307

Db 2309 YSLIE 2313

RESULT 6

US-09-105-537-37

; Sequence 37, Application US/09105537A

; Patent No. 6265202

; GENERAL INFORMATION:

; APPLICANT: Sherman, D. H.

; APPLICANT: Liu, H.

; APPLICANT: Xue, Y.

; APPLICANT: Zhao, L.

; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

; FILE REFERENCE: 600,438US1

; CURRENT APPLICATION NUMBER: US/09/105,537A

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 37

; LENGTH: 1346

; TYPE: PRT

; ORGANISM: Streptomyces venezuelae

US-09-105-537-37

Query Match 4.5%; Score 90.5; DB 4; Length 1346;

Best Local Similarity 21.5%; Pred. No. 1.1;

Matches 62; Conservative 20; Mismatches 87; Indels 119; Gaps 12;

Db 81 DGLMHEVNGLMERPDWETAIOKPLCSL-----PAGSG 113

Qy 854 DGGHRLTSLAE--WANGIALDMASILPATGALSPAVPDLPTFAFHRSYTWISPA 911

Db 114 NALAASLNHYAGEQVNTEDLLTNTCLLCRPVLSPMNLSTHTASGLRSTSVSLAMGF 173

Qy 912 EAPA-----HTASGSENAEISGLAMGP 933

Db 174 IADVDLESCKYRRIGEMRFTLCTPLRLAALRTYGRGLATLPYGR-----VGEKTPAS 225

Qy 934 GAE-DLDE-----GRRSAVLAMVRQAASVLRCDSPREVPVDRPLREIGFDSLAVDFR 987

Db 226 -----PVVVOGQPYDAHLV--LDEQVSSHQVQVPPDEFVLLALLSHSLASE 271

Qy 988 NRVRNLGLQLPPTVVFQHPFPVALAERISDELAERNAVAEPSP-----HEQAE 1039

Db 272 MEAAPMGR-----CAAGVMHLFYVRAGYSRAMLRFLAMEGRHMEY 314

Qy 1040 KAAPNAGARSGADTGAG-----AGMFRALFRQ--AVDDDRYGEF 1076

Db 1040 KAAPNAGARSGADTGAG-----AGMFRALFRQ--AVDDDRYGEF 1076

US-09-105-537-6

RESULT 7

US-09-105-537-6

; Sequence 6, Application US/09105537A

; Patent No. 6265202

; GENERAL INFORMATION:

; APPLICANT: Sherman, D. H.

; APPLICANT: Liu, H.

; APPLICANT: Xue, Y.

; APPLICANT: Zhao, L.

; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

; FILE REFERENCE: 600,438US1

; CURRENT APPLICATION NUMBER: US/09/105,537A

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 6

; LENGTH: 11877

; TYPE: PRT

; ORGANISM: Streptomyces venezuelae

US-09-105-537-6

Query Match 4.5%; Score 90.5; DB 4; Length 11877;
Best Local Similarity 21.5%; Pred. No. 33;
Matches 62; Conservative 20; Mismatches 87; Indels 119; Gaps 12;

QY 81 DGLMEVNVNLMRPRMERAIOKPLCLSL-----PAGSG 113
DB 11104 DGGHRTTSLA--WANGLADWASLDPATGALSPVDPDLFTYAFQHRSTWISPAFG 11161
QY 114 NALAASLNHYAGYEQVYNTEDLLTNCITLLCRPVLSPMNLISLTASGLRSFVSLANGF 173
DB 11162 EAPA-----HTASGREAVAEETGLANGP 11183
QY 174 IADVDLSDKYRRLGEMFTLGLFRLAALRTYRGLATLPYR-----VGEKTPAS----- 225
DB 11184 GAE-DLDE-----GRRSAVLAVMVAASVLRCDSPPEVVDRLPREIGFSLTAVDR 11237
QY 226 -----PVYVOOGPVDAHLVP--LEEQVPSHMOVVPDEDFVYLALHSHLASE 271
DB 11238 NRVRRLTGLDLPVTVFQHPTPVALARIISDELERNWAVAPSD-----HEQAEE 11289
QY 272 MEAPMGR-----CAAGVHLFTYRAGVSRAMLRLFLAMEKGRHMEY 314
DB 11290 KAAPAGARSGADTGA-----AGMFRALFRQ--AVEDDRYGE 11326

RESULT 8
US-08-936-135-2
Sequence 2, Application US/08936135
Patent No. 6054293

GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: He, Zhigang
APPLICANT: Chen, Hang
TITLE OF INVENTION: Semaphorin Receptors
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2588 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-936-135-2

Query Match 4.5%; Score 90; DB 3; Length 2588;

Best Local Similarity 23.8%; Pred. No. 3.4;
Matches 82; Conservative 34; Mismatches 135; Indels 94; Gaps 17;

QY 4 AGPRGVLPKPRVILVILNPRGKGKALQLFRSHVQPLLAELISFTLMLTERRHAREL 63

DB 1568 ARGPEHL-EARGILETYRPRGLANGALATHR-HISGLYG--LYLSGLYLEANGMETGL 1622
QY 64 VSEELGKMDALVMSGGL-----MEVNVGLMER-----PDMEIAIOK 103
DB 1623 LELEBLYCSGLVAL--GLALAPRTHRALGLPRHRTHRPRASNGLYASNI-EVALAS 1679
QY 104 P-LG-----SUPAGSG-NALAASLNHYAGYEQVYNTEDLLTNCITLLCRPVLSPMNLISLH 156
DB 1680 PELCTASAPASASGCLNALAAS-----NCYHISSEBRLYTHRGLYASAPSPHEGLNLE 1735
QY 157 TASGLRSFVSLANGFTIADVDLSDKYRRLGEMFTLGLFRLAALRTYRGLATLPYV 216
DB 1736 TRHGL-----YGLYTHRTHRA-----LEALATHRGGLYSPPTH 1770
QY 217 RVGEFTKPSPVY-----VOOGPVDAHLVPLEEQVPSHMOVVPDEDFVYLALHSHLASEM 272
DB 1771 RVALLIENASPSERTIRILEGLNSEGLPHEPRTHTYSG-----LYPHEAS-- 1816
QY 273 FAAPMGRCAAGVHML-FYVRAGVSRAMLRLFLAMEKGRHMEYEC 316
DB 1817 -----NCYSGLPHEGLYTRPGLYSERHISLY-----STHRPHEC 1850

RESULT 9
US-09-144-085-2
Sequence 2, Application US/09144085
Patent No. 6280999

GENERAL INFORMATION:
APPLICANT: Gustafsson, Claes
APPLICANT: Betlach, Mary C.
APPLICANT: Ashley, Gary
APPLICANT: Julien, Bryan
APPLICANT: Ziemann, Rainer
TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
FILE REFERENCE: 30062-20020.20
CURRENT APPLICATION NUMBER: US/09/144,085
CURRENT FILING DATE: 1998-08-31
EARLIER APPLICATION NUMBER: 09/010,809
EARLIER FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 6095
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-144-085-2

Query Match 4.4%; Score 88; DB 4; Length 6095;
Best Local Similarity 26.5%; Pred. No. 22;
Matches 75; Conservative 30; Mismatches 94; Indels 84; Gaps 16;

QY 1 MDPAGPRGVLPKPRVILVILNPRGKGKALQLFRSHVQPLLAELISFTLMLTERRN-- 58
DB 1427 LDPAG-----TVLV-----TGCTGEGGQAIH--LVRAVGVRHLVLTSSRGLE 1468
QY 59 --HARELYRS-EELGRMDALVMSGGLMEVNVNGLMERPDWETAIOKPLCSLPASGNA 115
DB 1469 APGARLEVOISLEEGA-ETVYVACADVSKKEEVARVLAGD-----AARPL----- 1513
QY 116 LAASLNHYAGYEQVYNTEDLLTNCITLLCRPVLS-----NMLISLTASGLRSFVSLA 170
DB 1514 --SAVHLHAG--VLDGVLTAQTAERLSVLAAPKVDGALHLELRELDLAFAVLFSSA 1568
QY 171 WFTIADVDLSDKYRRLGEMFTLGLFRLAALRTYRGLATLPVVRVGEKTPAS----- 225
DB 1569 AG-----TFGAGGOSNYAANTFLDALAAHRRGGGLAATSLA-WGFWQAGVGMT 1617
QY 226 -----PVYVOOG--PYDAHLV-PLLEQVPSH 248
DB 1618 AHLEPAELSRMRNGEVPMPVEBGLALIDALSPASVYPVH 1660

RESULT 10
US-09-134-001C-5553
Sequence 5553, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5553
LENGTH: 319
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5553

Query Match 4.3%; Score 87.5; DB 4; Length 319;
Best Local Similarity 23.7%; Pred. No. 0.24;
Matches 28; Conservative 26; Mismatches 47; Indels 17; Gaps 4;
QY 22 NPGGKALQLEFRSHVQPLAE--AEISFTMLTERNRHARELVSEELGRMDALVMS 79
DB 35 NGIGEVAKSLSSMCKHLSLQSEKGDIIKCKSIKENENSSD-----DVLFTLG 85
QY 80 GDGLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALASLNHAGYEQVYTNEDLTN 137
DB 86 GDGLMEHVNGVM-----QYQLNLPICGVIGGTENDTFTKTLQHPNF-KTASEQLLS 137

RESULT 11
US-09-320-878-4
Sequence 4, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
CURRENT FILING DATE: 1999-05-27
EARLIER APPLICATION NUMBER: CIP OF 09/141,908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846,247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/119,139
EARLIER FILING DATE: 1999-02-08
EARLIER APPLICATION NUMBER: 60/100,880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087,080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1346
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-320-878-4

Query Match 4.3%; Score 87.5; DB 3; Length 1346;
Best Local Similarity 21.2%; Pred. No. 2.3;
Matches 61; Conservative 21; Mismatches 87; Indels 119; Gaps 12;

QY 81 DGIMHEVNGLMERPDMETAIOKPLCSL-----PAGSG 113
DB 854 DGGHRLFTTSLAE--WANGIALDMASILPATGALSPAVPDLPTFYARQHRWSYWPAGG 911
QY 114 NALASLNHAGYEQVYTNEDLTNCTLLCPRVLSPNMLSLHNASGLRSVSLWAGF 173
DB 912 EAPF-----HTASGREAAETGLMGP 933
QY 174 IADVDLSDKYRGLGEMRFTLTGFLRLAALRTYGRATLPVGR-----VGKTPAS---- 225
DB 934 GAE-DLDEE-----GRSAVLAMVMRQAASVLRCDSEEVYVDPRLAIGDSLTADEF 987
QY 226 -----PVYVOGPDVAHLVP--LEEQVSHMOVYVDPDEFLVALLSHLASE 271
DB 988 NRVRNLGLQRPVYVEHPTPVALAERISDELERNMAVALEPSD-----HEQAE 1039
QY 272 MEAPMGR-----CAQVMHLFTYRAGVSRAMLRLFLAMEKGRHMEY 314
DB 1040 KAAPAGARSGADTGAG-----AGMFRALFRQ---AVEDRYGEF 1076

RESULT 12
US-09-335-409-2
Sequence 2, Application US/09335409
Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EROTHIOLONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1421
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-2

Query Match 4.2%; Score 85.5; DB 3; Length 1421;
Best Local Similarity 24.0%; Pred. No. 4.2; Indels 105; Gaps 19;
Matches 82; Conservative 34; Mismatches 121;
QY 25 GKGKALQLEFRSHVQPLAEAEISFTMLTERNRHAREL-VRSEELGRMDALVY-----M 78
DB 1099 GGVGLAIVQMAQH-----GAEVHATAGTPEKRAVLESLGVRYVSDSRDFADVRAWT 1153
QY 79 SGGLMEHVNGLMERPDMETAIOKPLCSLPAGSGNALASLN--HAGYEQVYTNEDL 135
DB 1154 GEGG-VDVVLNLSL-----SGELLDKSFNLSRSGREVEIGKRCY 1192
QY 136 TNCITLLCRPVLSPNMLSLHTASGLRSFVLSLWAGFIAVDLSDKYRGLGEMRFTLG 195
DB 1193 ADNOLGL-RPFL-RNL-----SFSL-----VDLRGMLEPRARYALLE 1229
QY 196 TFLRLAALRTYR-GRATLPVGRV--GKTPA-----SPVVOGPDVAHLVPLEQVPS 247
DB 1230 ELGLIAGVFTPPPIATLPFARVADAFRMAQOHGKLVLTIGDPEVQI-----RIFT 1284
QY 248 HMQVYP---DEDFVLVALLSHLASEMFAAPMGRCA-----GV 284
DB 1285 HAGAGPSTGDRD-----LLDRLAS--AAPARAALAEFLRTQVSQVLRTPKIVGA 1334
QY 285 MHLFVYRAGVSRAMLRLFLAMEKGRHMEYCPYLVVYVVA 326
DB 1335 EALF-TRLGMDSLAVALERNIASIKLSTTFLSTSPNIA 1375

```
RESULT 13
US-09-568-102-2
; Sequence 2, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-102-2

Query Match
Best Local Similarity 24.0%; Score 85.5; DB 4; Length 1421;
Matches 82; Conservative 34; Mismatches 121; Indels 105; Gaps 19;

OY 25 GKGKALQLFERSHQPLAEISFTLMLTERRNHREL-VRSEELGRMDALVY-----M 78
DB 1099 GGVGIAAVQMAOHV-----GAEVHATAGTPEKRAVLESIGRVYSDRSRDFVADYRAWT 1153
OY 79 SGDGIMHEVYVGLMERPDWETAIOKPLCSLPAGSNAALASLN---HYAGEVQTNEDLL 135
DB 1154 GGEQ-VDVYVLSL-----SGELIDKSFNLRSHGRFVELGKRDY 1192
OY 136 TNCITLLCRPYLSPNNLSLHTASGLRSFSVLSIANGFIADVDESCKYRRLGEMRFTLG 195
DB 1193 ADNQGTL-RPFL--RNL-----SFSL-----VDLRGMMLERPARVALL 1229
OY 196 TFLRLAALRYR-GRLATLPYGRV--GKTPA-----SPVYVQGPVDAHLVPLEEQVPS 247
DB 1230 ELGLIAGVFTPPPIATLPARVADAFRSMAQOHLKVLTLGDPEVQI-----RIP 1284
OY 248 HMQVVP--DEDFVLVLLSHSLASEMFAAPMGCA-----GV 284
DB 1285 HAGAGPSTGDRD-----LIDRLAS---AAPAAALAEAFRTQVSQVLRTEPIKVG 1334
OY 285 MHLFVYRAGVSRAMLRLFLAMEKGRHMEYECPLYVYVVA 326
DB 1335 EALF-TRIGMDSIMAVELNRIEASLKLSTTFSTSPNIA 1375

RESULT 14
US-09-567-969-2
; Sequence 2, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
```

```
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-567-969-2

Query Match
Best Local Similarity 24.0%; Score 85.5; DB 4; Length 1421;
Matches 82; Conservative 34; Mismatches 121; Indels 105; Gaps 19;

OY 25 GKGKALQLFERSHQPLAEISFTLMLTERRNHREL-VRSEELGRMDALVY-----M 78
DB 1099 GGVGIAAVQMAOHV-----GAEVHATAGTPEKRAVLESIGRVYSDRSRDFVADYRAWT 1153
OY 79 SGDGIMHEVYVGLMERPDWETAIOKPLCSLPAGSNAALASLN---HYAGEVQTNEDLL 135
DB 1154 GGEQ-VDVYVLSL-----SGELIDKSFNLRSHGRFVELGKRDY 1192
OY 136 TNCITLLCRPYLSPNNLSLHTASGLRSFSVLSIANGFIADVDESCKYRRLGEMRFTLG 195
DB 1193 ADNQGTL-RPFL--RNL-----SFSL-----VDLRGMMLERPARVALL 1229
OY 196 TFLRLAALRYR-GRLATLPYGRV--GKTPA-----SPVYVQGPVDAHLVPLEEQVPS 247
DB 1230 ELGLIAGVFTPPPIATLPARVADAFRSMAQOHLKVLTLGDPEVQI-----RIP 1284
OY 248 HMQVVP--DEDFVLVLLSHSLASEMFAAPMGCA-----GV 284
DB 1285 HAGAGPSTGDRD-----LIDRLAS---AAPAAALAEAFRTQVSQVLRTEPIKVG 1334
OY 285 MHLFVYRAGVSRAMLRLFLAMEKGRHMEYECPLYVYVVA 326
DB 1335 EALF-TRIGMDSIMAVELNRIEASLKLSTTFSTSPNIA 1375

RESULT 15
US-09-568-480-2
; Sequence 2, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-480-2

Query Match
Best Local Similarity 24.0%; Score 85.5; DB 4; Length 1421;
Matches 82; Conservative 34; Mismatches 121; Indels 105; Gaps 19;

OY 25 GKGKALQLFERSHQPLAEISFTLMLTERRNHREL-VRSEELGRMDALVY-----M 78
DB 1099 GGVGIAAVQMAOHV-----GAEVHATAGTPEKRAVLESIGRVYSDRSRDFVADYRAWT 1153
OY 79 SGDGIMHEVYVGLMERPDWETAIOKPLCSLPAGSNAALASLN---HYAGEVQTNEDLL 135
DB 1154 GGEQ-VDVYVLSL-----SGELIDKSFNLRSHGRFVELGKRDY 1192
OY 136 TNCITLLCRPYLSPNNLSLHTASGLRSFSVLSIANGFIADVDESCKYRRLGEMRFTLG 195
DB 1193 ADNQGTL-RPFL--RNL-----SFSL-----VDLRGMMLERPARVALL 1229
OY 196 TFLRLAALRYR-GRLATLPYGRV--GKTPA-----SPVYVQGPVDAHLVPLEEQVPS 247
DB 1230 ELGLIAGVFTPPPIATLPARVADAFRSMAQOHLKVLTLGDPEVQI-----RIP 1284
OY 248 HMQVVP--DEDFVLVLLSHSLASEMFAAPMGCA-----GV 284
DB 1285 HAGAGPSTGDRD-----LIDRLAS---AAPAAALAEAFRTQVSQVLRTEPIKVG 1334
OY 285 MHLFVYRAGVSRAMLRLFLAMEKGRHMEYECPLYVYVVA 326
DB 1335 EALF-TRIGMDSIMAVELNRIEASLKLSTTFSTSPNIA 1375
```

```

Db      1193  ADNOLGL-RPFL--RNL-----SFSL-----VDLRGMMLERPARVRALE 1229
QY      1396  TFLRLAALRTYR-GRLATLPYGRV--GFKTPA-----SPVYVQOGPVDAHLVPLEEQVPS 247
Db      1230  ELLGLIAGVFTPPPIATLPIARVADAFRSMAQOHGKLVTLGDPEVOI-----RIP 1284
QY      248  HMOYVP---DEDFVLVLALLSHSLASEMFAPMGRCA-----GV 284
Db      1285  HAGGPFSTGDRD-----LIDRLAS--AAPAARAALAEFLRTQVSQVLTPEIKYGA 1334
QY      285  MHLFYVRAGVSRAMLRLFLAMEKGRHMEYECPIVYVYVYA 326
Db      1335  EALF-TRLGMDSLMAVELRNRIEASLKLSTFISTSPNTA 1375

```

Search completed: July 12, 2003, 18:10:05
 Job time : 21 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2003, 18:00:22 ; Search time 46 Seconds

(without alignments)
1720.045 Million cell updates/sec

Title: US-09-784-810A-2

Perfect score: 2015

Sequence: 1 MDPAGGPRGVLPKRCRVLYL.....CPEPPPSMKPQGMPPPEPL 384

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1955	97.0	384	4	Q96HV8
2	1945	96.5	398	4	Q96GK1
3	1590	78.9	329	6	Q9N0A5
4	1585.5	78.7	381	11	Q912N3
5	1585.5	78.7	388	11	Q88886
6	1582.5	78.5	504	11	Q88885
7	1555.5	77.2	383	11	Q91V26
8	1454	72.2	290	4	Q9BTG7
9	856.5	42.5	654	4	Q9H0Q2
10	856.5	42.5	654	4	Q9BRN1
11	590	29.3	384	11	Q9DAD1
12	550	27.3	641	5	Q9VYV8
13	494	24.5	661	5	Q8T0C1
14	491.5	24.4	907	5	Q9VZW0
15	438	21.7	1240	10	Q65419
16	403	20.0	624	5	Q8T269

17	372.5	18.5	473	5	Q18425	Q18425 caenorhabdi
18	369	18.3	458	3	Q14159	Q14159 schizosach
19	363	18.0	354	4	Q9NMU7	Q9NMU7 homo sapien
20	355.5	17.6	687	3	Q06147	Q06147 saccharomyc
21	337	16.7	624	3	Q12246	Q12246 saccharomyc
22	271	13.4	763	10	Q9LRB0	Q9LRB0 arabidopsis
23	263.5	13.1	481	4	Q9BYB3	Q9BYB3 homo sapien
24	263.5	13.1	537	4	Q8TCT0	Q8TCT0 homo sapien
25	253	12.6	757	10	Q94HT9	Q94HT9 oryza sativ
26	249	12.4	732	10	Q9FHL3	Q9FHL3 arabidopsis
27	233.5	11.6	549	5	Q9T211	Q9T211 caenorhabdi
28	233	11.6	487	5	Q95S15	Q95S15 drosophila
29	233	11.6	596	5	Q9YNA6	Q9YNA6 drosophila
30	223	11.1	326	4	Q9UGC5	Q9UGC5 homo sapien
31	205	10.2	422	4	Q9NP48	Q9NP48 homo sapien
32	198	9.8	421	11	Q9ESV4	Q9ESV4 mus musculu
33	176.5	8.8	533	10	Q9LU45	Q9LU45 arabidopsis
34	164.5	8.2	406	5	Q9VLI0	Q9VLI0 drosophila
35	148	7.3	173	11	Q9D087	Q9D087 mus musculu
36	141.5	7.0	306	16	Q927N6	Q927N6 listeria in
37	140.5	7.0	338	10	Q94C71	Q94C71 arabidopsis
38	133	6.6	302	16	Q9CEN8	Q9CEN8 lactococcus
39	132.5	6.6	309	16	Q34799	Q34799 bacillus su
40	131	6.5	586	10	Q949C3	Q949C3 oryza sativ
41	130.5	6.5	303	16	Q31502	Q31502 bacillus su
42	129.5	6.4	306	16	Q8Y497	Q8Y497 listeria mo
43	125	6.2	295	16	Q9KBH4	Q9KBH4 bacillus ha
44	123	6.1	291	16	Q8R615	Q8R615 thermocane
45	122.5	6.1	311	16	Q9K000	Q9K000 bacillus ha

ALIGNMENTS

RESULT 1

Q96HV8 PRELIMINARY: PRT: 384 AA.

AC Q96HV8: 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)

DE Unknown (protein for MGC:15041).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RA NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RC TISSUE=SKIN;

RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL, BC008040, AA08040.1, -.

DR InterPro: IPR001395; Aldo/Ket_red.

DR InterPro: IPR001206; DAGKC.

DR InterPro: IPR003622; DAG_kin_cat.

DR Pfam: PF00781; DAGKC; 1.

DR ProDom: PD005043; DAG_kin_cat; 1.

DR PROSITE: PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.

SQ SEQUENCE 384 AA; 42474 MW; F82999FF306113B0 CRC64;

Query Match 97.0%; Score 1955; DB 4; Length 384;
Best Local Similarity 97.4%; Pred. No. 1.6e-164;
Matches 374; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY	1	MDPAGGPRGVLPKRCRVLYLNP	RGKKAQLQFRSHVQPLLA	EAIEISFTLMLTERRNHA	60
DB	1	MDPAGGPRGVLPKRCRVLYLNP	RGKKAQLQFRSHVQPLLA	EAIEISFTLMLTERRNHA	60
QY	61	RELVAEELGRWDALVYNSG	GLMEHYVNGLMERDWEA	IAIQPLCSLPAGSGNLA	120
DB	61	RELVAEELGRWDALVYNSG	GLMEHYVNGLMERDWEA	IAIQPLCSLPAGSGNLA	120
QY	121	NHYAGYEQVTNEDLLTNC	TLLCRPVLPSPMNLST	LHTFASGLRFSVLAMGFI	180
DB	121	NHYAGYEQVTNEDLLTNC	TLLCRPVLPSPMNLST	LHTFASGLRFSVLAMGFI	180

DB 121 NHYAGYEQVTNEDLTNCTLLCRRLSPMNLSTLTASGLRFVSLAMGFADVDLE 180
QY 181 SDXYRRLGEMRFTLGTFLRLAALRTYRGLATLPVGRVGFKTPASPVVVOGSPVDAHLP 240
DB 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGLATLPVGRVSKTPASPVVVOGSPVDAHLP 240
QY 241 LEEOVPSHMOVPEDEDEVLVALLSHSLASEMFAAPMGRCAAGVMHLFTYRAGVSRAML 300
DB 241 LEEVPSHMTVPDEDEFLVALLSHSLASEMFAAPMGRCAAGVMHLFTYRAGVSRAML 300
QY 301 RLFLAMKGRHMEYECPLYVYVVAFRLEPKDKGVAVDGLAMVSEAVOGVHPNFTM 360
DB 301 RLFLAMKGRHMEYECPLYVYVVAFRLEPKDKGVAVDGLAMVSEAVOGVHPNFTM 360
QY 361 MWSGCVPEPPSWKPOQMPPEEPL 384
DB 361 MWSGCVPEPPSWKPOQMPPEEPL 384

RESULT 2

Q96GK1 PRELIMINARY; PRT; 398 AA.
ID 096GK1
AC 096GK1
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:15040).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC009419; AAH09419.1; -
DR InterPro: IPR001395; Aldo/Ket_red.
DR InterPro: IPR001206; DAGC.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGC; 1.
DR Prodom: PD005043; DAG_kin_cat; 1.
DR PROSITE: PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
SQ SEQUENCE 398 AA; 43944 MW; 7FC9B26C030E560 CRC64;

Query Match 96.5%; Score 1945; DB 4; Length 398;

Best Local Similarity 94.2%; Pred. No. 1,3e-163; Indels 14; Gaps 1;
Matches 375; Conservative 2; Mismatches 7;

QY 1 MDP-----AGPRGVLPKRCRVLYLNLNPRGKGKALQFRSHVOPLLAEAE 46
DB 1 MDPVVGCGKGLFGFPGSAGPRGVLPKRCRVLYLNLNPRGKGKALQFRSHVOPLLAEAE 60
QY 47 ISFTLMLTERRNHARELYRSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIQKPLC 106
DB 61 ISFTLMLTERRNHARELYRSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIQKPLC 120
QY 107 SLPAGSGNALAASLIMHYAGYEQVTNEDLTNCTLLCRPVLSPMNLSTLTASGLRSFSV 166
DB 121 SLPAGSGNALAASLIMHYAGYEQVTNEDLTNCTLLCRPVLSPMNLSTLTASGLRSFSV 180
QY 167 LSLAMGFADVDLESDKYRRLGEMRFTLGTFLRLAALRTYRGLATLPVGRVGFTPASP 226
DB 181 LSLAMGFADVDLESDKYRRLGEMRFTLGTFLRLAALRTYRGLATLPVGRVGFTPASP 240
QY 227 VVVQGPVDAHLVPLEEOVPSHMOVPEDEDEVLVALLSHSLASEMFAAPMGRCAAGVMH 286
DB 241 VVVQGPVDAHLVPLEEOVPSHMTVPDEDEFLVALLSHSLASEMFAAPMGRCAAGVMH 300
QY 287 LFTYRAGVSRAMLRLFLAMKGRHMEYECPLYVYVVAFRLEPKDKGVAVDGLAMV 346
DB 301 LFTYRAGVSRAMLRLFLAMKGRHMEYECPLYVYVVAFRLEPKDKGVAVDGLAMV 360

QY 347 SEAVOGVHPNFTMWSGCVPEPPSWKPOQMPPEEPL 384
DB 361 SEAVOGVHPNFTMWSGCVPEPPSWKPOQMPPEEPL 398

RESULT 3

Q9NOA5 PRELIMINARY; PRT; 329 AA.
ID 09NOA5
AC 09NOA5
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Unnamed protein product.
OS Macaca fascicularis (Crib eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_Taxid=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046025; BAB01607.1; -
DR InterPro: IPR001395; Aldo/Ket_red.
DR InterPro: IPR001206; DAGC.
DR InterPro: IPR003622; DAG_kin_cat.
DR InterPro: IPR00781; DAGC; 1.
DR Pfam: PF00781; DAGC; 1.
DR Prodom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00046; DAGC; 1.
DR PROSITE: PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
SQ SEQUENCE 329 AA; 36393 MW; 7F430ABD0CA5FDC7 CRC64;

Query Match 78.9%; Score 1590; DB 6; Length 329;

Best Local Similarity 95.3%; Pred. No. 2.4e-132; Indels 0; Gaps 0;
Matches 307; Conservative 4; Mismatches 11;

QY 1 MDPAGPRGVLPKRCRVLYLNLNPRGKGKALQFRSHVOPLLAEAEISFTLMLTERRNHA 60
DB 1 MDPAGPRGVLPKRCRVLYLNLNPRGKGKALQFRSHVOPLLAEAEISFTLMLTERRNHA 60
QY 61 RELVRSSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIQKPLCSLPAGSGNALAASL 120
DB 61 RELVRSSEELGRMDALVYMSGDGLMHEVYNGIMERPDMETAIQKPLCSLPAGSGNALAASL 120
QY 121 NHYAGYEQVTNEDLTNCTLLCRPVLSPMNLSTLTASGLRSFSVSLAMGFADVDLE 180
DB 121 NHYAGYEQVTNEDLTNCTLLCRPVLSPMNLSTLTASGLRSFSVSLAMGFADVDLE 180
QY 181 SDXYRRLGEMRFTLGTFLRLAALRTYRGLATLPVGRVGFKTPASPVVVOGSPVDAHLP 240
DB 181 SEKYRRLGEMRFTLGTFLRLAALRTYRGLATLPVGRVSKTPASPVVVOGSPVDAHLP 240
QY 241 LEEOVPSHMOVPEDEDEVLVALLSHSLASEMFAAPMGRCAAGVMHLFTYRAGVSRAML 300
DB 241 LEEVPSHMTVPDEDEFLVALLSHSLASEMFAAPMGRCAAGVMHLFTYRAGVSRAML 300
QY 301 RLFLAMKGRHMEYECPLYVYVVAFRLEPKDKGVAVDGLAMVSEAVOGVHPNFTM 360
DB 301 RLFLAMKGRHMEYECPLYVYVVAFRLEPKDKGVAVDGLAMVSEAVOGVHPNFTM 360

RESULT 4

Q91ZN3 PRELIMINARY; PRT; 381 AA.
ID 091ZN3
AC 091ZN3
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sphingosine kinase 1a.

GN SPHK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Thompson D., Pyne S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF415213; AL07499.1; -
DR MGI: 1316649; Spkl1.
DR InterPro: IPR001206; DAGKC.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKC; 1.
DR Prodom: PD005043; DAG_kin_cat; 1.
KW Kinase.
SQ SEQUENCE 381 AA; 42344 MW; 2F2C1E10D59EB129 CRC64;

Query Match 78.7%; Score 1585.5; DB 11; Length 381;
Best Local Similarity 79.8%; Pred. No. 7.3e-132;
Matches 301; Conservative 34; Mismatches 41; Indels 1; Gaps 1;

QY 7 PGVLPRPCRYLVILNPRGKGKALQIFRSHVOPPLAEAEISFTLMTERRNHARELYRS 66
DB 6 PGLLPKRCRYLVILNPGGKGKALQIFSRVOPPLEAEITFKLITERKNHARELYCA 65
QY 67 BELGMDALVYMSGDLGHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGY 126
DB 66 BELGMDALVYMSGDLGHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASVNHYAGY 125
QY 127 EOVYNEDLLTNCITLLCRPVLSPNMLSLHTASGLRSVSLANGFTADVDLESCKYR 186
DB 126 EOVYNEDLLTNCITLLCRPVLSPNMLSLHTASGLRSVSLANGFTADVDLESCKYR 185
QY 187 LGEMFTLGTFLRLAALRTYRGRLATLPVGRVGFTRPASPYYVOGQPVDAHLVPLEEQVP 246
DB 186 LGEIFRTVGTFFRLASLRYOGQLAYLPVGTVASRRPAS-TLVQGPVDTHLVPLEEVP 244
QY 247 SHMOVPPDEDFVLVYLLALSHSLASMFAPMGRCAGVYMHLYFYVAGVSRAMLRLFLAM 306
DB 245 SHMTVPPQDPFLVYLLHTHLSSELFAPMGRCAAGVYMHLYFYVAGVSRAMLRLFLAM 304
QY 307 EKGRIEYECPLVYVYVPAFAFLPEKDKGFVAVDGEIYMSAVNOGVHPNYFMVYSCV 366
DB 305 QKGRIMELDCPYLVAVPVPAFAFLPEKRSRGVFSVDGELMCEAVNOGVHPNYFMVYSCV 364
QY 367 EPPPSWKPOQMPPEEP 383
DB 365 DAPSGRDSRGRGPPPEEP 381

RESULT 5
088886 PRELIMINARY; PRT; 388 AA.
AC 088886;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Sphingosine kinase.
GN SPHK1 OR SPHK1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-98395082; PubMed-9726979;
RA Kohama T., Olivera A., Edsall L., Nagiec M.M., Dickson R., Spiegel S.;
RT "Molecular cloning and functional characterization of murine
RT sphingosine kinase."
RT J. Biol. Chem. 273:23722-23728(1998).
DR EMBL: AF068749; AAC61698.1; -
DR MGI: 1316649; Spkl1.

DR InterPro: IPR001206; DAGKC.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKC; 1.
DR Prodom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00046; DAGKC; 1.
KW Kinase.
SQ SEQUENCE 388 AA; 43254 MW; 01EB032322542CAD CRC64;

Query Match 78.7%; Score 1585.5; DB 11; Length 388;
Best Local Similarity 79.8%; Pred. No. 7.5e-132;
Matches 301; Conservative 34; Mismatches 41; Indels 1; Gaps 1;

QY 7 PGVLPRPCRYLVILNPRGKGKALQIFRSHVOPPLAEAEISFTLMTERRNHARELYRS 66
DB 13 PGLLPKRCRYLVILNPGGKGKALQIFSRVOPPLEAEITFKLITERKNHARELYCA 72
QY 67 BELGMDALVYMSGDLGHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGY 126
DB 73 BELGMDALVYMSGDLGHEVYNGLMERPDWETAIOKPLCSLPAGSGNALAASVNHYAGY 132
QY 127 EOVYNEDLLTNCITLLCRPVLSPNMLSLHTASGLRSVSLANGFTADVDLESCKYR 186
DB 133 EOVYNEDLLTNCITLLCRPVLSPNMLSLHTASGLRSVSLANGFTADVDLESCKYR 192
QY 187 LGEMFTLGTFLRLAALRTYRGRLATLPVGRVGFTRPASPYYVOGQPVDAHLVPLEEQVP 246
DB 193 LGEIFRTVGTFFRLASLRYOGQLAYLPVGTVASRRPAS-TLVQGPVDTHLVPLEEVP 251
QY 247 SHMOVPPDEDFVLVYLLALSHSLASMFAPMGRCAGVYMHLYFYVAGVSRAMLRLFLAM 306
DB 252 SHMTVPPQDPFLVYLLHTHLSSELFAPMGRCAAGVYMHLYFYVAGVSRAMLRLFLAM 311
QY 307 EKGRIEYECPLVYVYVPAFAFLPEKDKGFVAVDGEIYMSAVNOGVHPNYFMVYSCV 366
DB 312 QKGRIMELDCPYLVAVPVPAFAFLPEKRSRGVFSVDGELMCEAVNOGVHPNYFMVYSCV 371
QY 367 EPPPSWKPOQMPPEEP 383
DB 372 DAPSGRDSRGRGPPPEEP 388

RESULT 6
088885 PRELIMINARY; PRT; 504 AA.
AC 088885;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Sphingosine kinase (Fragment).
GN SPHK1 OR SPHK1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-98395082; PubMed-9726979;
RA Kohama T., Olivera A., Edsall L., Nagiec M.M., Dickson R., Spiegel S.;
RT "Molecular cloning and functional characterization of murine
RT sphingosine kinase."
RT J. Biol. Chem. 273:23722-23728(1998).
DR EMBL: AF068748; AAC61697.1; -
DR MGI: 1316649; Spkl1.
DR InterPro: IPR001206; DAGKC.
DR InterPro: IPR003622; DAG_kin_cat.
DR Pfam: PF00781; DAGKC; 1.
DR Prodom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00046; DAGKC; 1.
KW Kinase.
FT NON_TER
SQ SEQUENCE 504 AA; 55119 MW; 2AAFEBA72027F509 CRC64;

Query Match 78.5%; Score 1582.5; DB 11; Length 504;

Best Local Similarity 79.6%; Pred. No. 1.9e-131;
Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

```
OY 7 PRGVLPKPRVVLINPRGKGKALOLFERSHVQPLLAEMISFTLMLERRRHAREIVRS 66
DB 129 PRLCLPRPCRVVLINPRGKGKALOLFOSRVOPLLEAEITFKILLTERKHAHELVA 188
OY 67 EEELGRDALVYMSGDGLMHEVYVNGLMERPDWETAIOKPLCSLPAGSGNALAASLNHYAGY 126
DB 189 EEELGHMDALVYMSGDGLMHEVYVNGLMERPDWETAIOKPLCSLPAGSGNALAASVNHVYAGY 248
OY 127 EGVYNEEDLLNCTLLCRLPVLSPMNLISHTASGLRSFVSLAMGFIADVDESCKYR 186
DB 249 EGVYNEEDLLNCTLLCRLPVLSPMNLISHTASGLRLYSVLISWGFVADVDESCKYR 308
OY 187 LGEMRFLGTFPLRLAALRTYRGRLATLPVGRVGFKTPTASPVVVOGVPVDAHLVPLEEQVP 246
DB 309 LGELRTVGFTEFRSLAKITVQGLATLPVGTVAASKRPAS-TLYQKGPVDTHVPLEEPPV 367
OY 247 SHMOVDEDEFLVTLALSHSLASEMFAPMGRCAAGVMHLFYVRAGVSRAMLRLFLAM 306
DB 368 SHMTVVEQDFLVLTLVLLHTLSSELPAAPMGRCAAGVMHLFYVRAGVSRAMLRLFLAM 427
OY 307 EKGHRMECEPTLVYVVAFRLEPKDKGVFAVDGELMYSVAOVGVHNYFMVYSGCV 366
DB 428 QKGMELDCPYLVHVPVAFRLERPSQGVFSVDGELMYSVAOVGVHNYFMVYSGCV 487
OY 367 EPPPSMKPQOMPPEEP 383
DB 488 DAPSGRDSRGRPPPEEP 504
```

RESULT 7
O91V26 PRELIMINARY; PRT; 383 AA.

```
AC 091V26:
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 20, Last annotation update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Sphingosine kinase 1f (Sphingosine kinase 1c)
DE Sphingosine kinase 1d (Sphingosine kinase 1e).
GN SPHK1F OR SPHK1A OR SPHK1C OR SPHK1D OR SPHK1E.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
OX [1]
RN SEQUENCE FROM N.A.
RA Imamura T., Ohgawa J., Ito S., Ogawa T., Hattori N., Tanaka S.,
RA Shiohara K.;
RT "cpg island of rat sphingosine kinase-1 gene: tissue-dependent DNA
RT methylation status and multiple alternative first exons.";
RL Genomics 78:117-125(2001).
RL EMBL; AB049575; BAB62324.1; -
DR EMBL; AB049571; BAB62320.1; -
DR EMBL; AB049572; BAB62321.1; -
DR EMBL; AB049573; BAB62322.1; -
DR EMBL; AB049574; BAB62323.1; -
DR EMBL; AB049575; BAB62323.1; -
DR InterPro: IPR001206; DAGKc.
DR InterPro: IPR001206; DAGKc.
DR InterPro: IPR001206; DAGKc.
DR Pfam; PF00781; DAGKc; 1.
DR ProDom; PD005043; DAG_kin_cat; 1.
KW KINASE.
SQ SEQUENCE 383 AA; 42418 MW; CAAD5817BF4B9507 CRC64;
```

Query Match 77.2%; Score 155.5; DB 11; Length 383;
Best Local Similarity 76.3%; Pred. No. 3.3e-129;
Matches 293; Conservative 44; Mismatches 46; Indels 1; Gaps 1;

```
OY 1 MDPAGGPRGVLPKRCRVVLINPRGKGKALOLFERSHVQPLLAEMISFTLMLERRRH 60
DB 1 MDPADCRGLLPKRCRVVLINPRGKGKALOLFOSRVOPLLEAEITFKILLTERKHA 60
```

```
OY 61 RELVSRBELGRMDALVYMSGDGLMHEVYVNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
DB 61 RELVCAEELGHMDALVYMSGDGLMHEVYVNGLMERPDWETAIOKPLCSLPAGSGNALAASL 120
OY 121 NHVAYBOVYNEDLLNCTLLCRLPVLSPMNLISHTASGLRSFVSLAMGFIADVDESCKYR 180
DB 121 NYVAGHOVYNEDLLNCTLLCRLPVLSPMNLISHTASGLRSFVSLAMGFIADVDESCKYR 180
OY 181 SDKTRRLGEMRFLGTFPLRLAALRTYRGRLATLPVGRVGFKTPTASPVVVOGVPVDAHLV 240
DB 181 SEKRSJGELRTVGFTEFRSLAKITVQGLATLPVGTVAASKRPAS-TLYQKGPVDTHVPLEEPPV 239
OY 241 LEEQVPSHMOVDEDEFLVTLALSHSLASEMFAPMGRCAAGVMHLFYVRAGVSRAMLRLFLAM 300
DB 240 LEEQVPSHMOVDEDEFLVTLALSHSLASEMFAPMGRCAAGVMHLFYVRAGVSRAMLRLFLAM 300
OY 301 RLFLAMEGRMECEPTLVYVVAFRLEPKDKGVFAVDGELMYSVAOVGVHNYFMVYSGCV 360
DB 300 RLFLAMQKGMHMDLDCPYLVHVPVAFRLERPSQGVFSVDGELMYSVAOVGVHNYFMVYSGCV 359
OY 361 MYSGCEPPEPSMKPQOMPPEEP 384
DB 360 MYSGSDSPSGRDSRGRPPPEEP 383
```

RESULT 8
O9BTG7 PRELIMINARY; PRT; 290 AA.

```
AC 09BTG7:
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to sphingosine kinase 1 (fragment).
DE Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN SEQUENCE FROM N.A.
RA TISSUE-SKIN;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC004112; AAH04112.1; -
RW KINASE.
KW NON_TER
FT SEQUENCE 290 AA; 31998 MW; A3B7A219DB52C520 CRC64;
```

Query Match 72.2%; Score 1454; DB 4; Length 290;
Best Local Similarity 96.9%; Pred. No. 2.1e-120;
Matches 278; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```
OY 98 ETATQKPLCSLPAGSGNALAASLNHYAGYBOVYNEDLLNCTLLCRLPVLSPMNLISHT 157
DB 4 ETATQKPLCSLPAGSGNALAASLNHYAGYBOVYNEDLLNCTLLCRLPVLSPMNLISHT 63
OY 158 ASGLRSFVSLAMGFIADVDESCKYRRLGEMRFLGTFPLRLAALRTYRGRLATLPVGR 217
DB 64 ASGLRLRSFVSLAMGFIADVDESCKYRRLGEMRFLGTFPLRLAALRTYRGRLATLPVGR 123
OY 218 VGRFTPTASPVVVOGVPVDAHLVPLEEQVPVSHMOVDEDEFLVTLALSHSLASEMFAPM 277
DB 218 VGRFTPTASPVVVOGVPVDAHLVPLEEQVPVSHMOVDEDEFLVTLALSHSLASEMFAPM 183
OY 124 VGSRTPTASPVVVOGVPVDAHLVPLEEQVPVSHMOVDEDEFLVTLALSHSLASEMFAPM 183
DB 124 VGSRTPTASPVVVOGVPVDAHLVPLEEQVPVSHMOVDEDEFLVTLALSHSLASEMFAPM 183
OY 278 GRCAAGVNHLYVYVAGVSRAMLRLFLAMEGRMECEPTLVYVVAFRLEPKDKGV 337
DB 184 GRCAAGVNHLYVYVAGVSRAMLRLFLAMEGRMECEPTLVYVVAFRLEPKDKGV 243
OY 338 FAVDGLMYSVAOVGVHNYFMVYSGCEPPEPSMKPQOMPPEEP 384
DB 244 FAVDGLMYSVAOVGVHNYFMVYSGCEPPEPSMKPQOMPPEEP 290
```

RESULT 9

Q9H002 PRELIMINARY; PRT; 654 AA.
 ID Q9H002
 AC Q9H002
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 69.2 kDa protein.
 GN DKF2564M0682.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRIN;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansgorge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler R., Strack N.,
 RA Mewes H.W., Othenwaelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wandt R., Korn B., Klein M., Poustka A.,
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 RL EMBL: AL136701; CAB6636.1;
 DR InterPro: IPR001206; DAGKC;
 DR InterPro: IPR003622; DAG_kin_cat.
 DR Pfam: PF00781; DAGKC; 1.
 DR ProDom: PD005043; DAG_kin_cat; 1.
 DR SMART: SM00046; DAGKC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 654 AA; 69207 MW; E9DF12E2AC0CF0B CRC64;

Query Match 42.5%; Score 856.5; DB 4; Length 654;
 Best Local Similarity 38.6%; Pred. No. 4e-67;
 Matches 189; Conservative 63; Mismatches 111; Indels 127; Gaps 6;

QY 10 VLPRCRLVILNPRGGKALQLFRSHVQPLAEAFSLTLMTERRNHARELYRSEL 69
 DB 176 LIPRRRLILVNPFGGKGLAMQCKNHYLPISAGSLFNIQTERONHARELYQGLSL 235
 QY 70 GRMDALVYMSGDGLMHEVYNGIMERPDMETAIQKPLCSLPAGSGNALASLNHYAGYBOY 129
 DB 236 SEMDGIIVYSGDGLIHEVYNGILDRPDWEAVKMPVGIIPCGSNALAGAVYQHGFPBA 295
 QY 130 TNEDLITCTLLCPVLSPNMLSLHTASGLRSYSVLSLNGFIADVDLESCKRRLGE 189
 DB 296 LGDLILLNCSTLRCGGGHPDLDSVTLASGRCSFSLVAMGFVSDVDIOSERRFALGS 355
 QY 190 MRFITGFLRLAALRTYGRGLATLPV----- 215
 DB 356 ARFTIGVLGLATLTHTYGRGLSYLPATVPASPPRAHSLPRAKSELTLTPDPAPMAHSP 415
 QY 216 ----- 215
 DB 416 LHRVSVDLPRLPOPALASPGSPERPLPILSLNCGGPELAGDMGAGDAPLSPDLLSSPP 475
 QY 216 --GRVGEKTPAS--PVVVQO-----GVDNALVPLEEOVPSHMOYVP 253
 DB 476 GSPKALHSPVSEGAIVIPSSGILPLPPDARVGAAGTGGPPHLLPPLGTPLPDPW-VTL 534
 QY 254 DEDFLVIALHSHSLASEMPAARPGRCAGVNHLYVYAGVSRAMLRLFLAMEGRHME 313
 DB 535 EGDFFVLMALISPSHIGADLVAAAPARFDDGLVHLCAVWSGISRAALLRLFLAMEGRSHFS 594
 QY 314 YECPLYLVVYVAFLEPKDGKGFVAVDGLMVSVAOVGOVHPNTFMVYSGCVPEPPSMK 373
 DB 595 LGGPQLGYAAARAFLEPLTPRGVLTVDGEQVEYERPLDAQNHPIGILTLLTG----PQCG- 649
 QY 374 PQQMPPEEP 383
 DB 650 -----PGRGP 654

RESULT 10
 ID Q9BRN1 PRELIMINARY; PRT; 654 AA.
 AC Q9BRN1
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 69.2 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LYMPH;
 RA Strausberg R.;
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
 RL EMBL: BC006161; AAH06161.1;
 DR InterPro: IPR001206; DAGKC;
 DR InterPro: IPR003622; DAG_kin_cat.
 DR Pfam: PF00781; DAGKC; 1.
 DR ProDom: PD005043; DAG_kin_cat; 1.
 DR SMART: SM00046; DAGKC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 654 AA; 69217 MW; F73FCEC930DA50F CRC64;

Query Match 42.5%; Score 856.5; DB 4; Length 654;
 Best Local Similarity 38.6%; Pred. No. 4e-67;
 Matches 189; Conservative 63; Mismatches 111; Indels 127; Gaps 6;

QY 10 VLPRCRLVILNPRGGKALQLFRSHVQPLAEAFSLTLMTERRNHARELYRSEL 69
 DB 176 LIPRRRLILVNPFGGKGLAMQCKNHYLPISAGSLFNIQTERONHARELYQGLSL 235
 QY 70 GRMDALVYMSGDGLMHEVYNGIMERPDMETAIQKPLCSLPAGSGNALASLNHYAGYBOY 129
 DB 236 SEMDGIIVYSGDGLIHEVYNGILDRPDWEAVKMPVGIIPCGSNALAGAVYQHGFPBA 295
 QY 130 TNEDLITCTLLCPVLSPNMLSLHTASGLRSYSVLSLNGFIADVDLESCKRRLGE 189
 DB 296 LGDLILLNCSTLRCGGGHPDLDSVTLASGRCSFSLVAMGFVSDVDIOSERRFALGS 355
 QY 190 MRFITGFLRLAALRTYGRGLATLPV----- 215
 DB 356 ARFTIGVLGLATLTHTYGRGLSYLPATVPASPPRAHSLPRAKSELTLTPDPAPMAHSP 415
 QY 216 ----- 215
 DB 416 LHRVSVDLPRLPOPALASPGSPERPLPILSLNCGGPELAGDMGAGDAPLSPDLLSSPP 475
 QY 216 --GRVGEKTPAS--PVVVQO-----GVDNALVPLEEOVPSHMOYVP 253
 DB 476 GSPKALHSPVSEGAIVIPSSGILPLPPDARVGAAGTGGPPHLLPPLGTPLPDPW-VTL 534
 QY 254 DEDFLVIALHSHSLASEMPAARPGRCAGVNHLYVYAGVSRAMLRLFLAMEGRHME 313
 DB 535 EGDFFVLMALISPSHIGADLVAAAPARFDDGLVHLCAVWSGISRAALLRLFLAMEGRSHFS 594
 QY 314 YECPLYLVVYVAFLEPKDGKGFVAVDGLMVSVAOVGOVHPNTFMVYSGCVPEPPSMK 373
 DB 595 LGGPQLGYAAARAFLEPLTPRGVLTVDGEQVEYERPLDAQNHPIGILTLLTG----PQCG- 649
 QY 374 PQQMPPEEP 383
 DB 650 -----PGRGP 654

RESULT 11
 ID Q9D4D1 PRELIMINARY; PRT; 384 AA.
 AC Q9D4D1
 DT 01-JUN-2001 (TREMBLrel. 17, Created)


```

Db 304 PREPR-----IHAALTCMAGKSTPMDDYVRVELARDKHFVYSLSYGWLADIDIES 359
OY 182 DYYRRLGEMRFTLGFELRLAALRTYRGRLATPVGRVGRKTPASVY----- 227
Db 360 ERLSRIGAKRFTLMAIKRLIGRSTYKGRVSYL-LCK-GKRP--PVEAARELPASTAAG 415
OY 228 VVOQGPVDA--HLVPLEE----- 243
Db 416 IRSSJPLNAGEFHDPPEEEGAVLDGDFADALSUDSVYRQHADSMSHSMSRTATYS 475
OY 244 -----OVV-----SHMOVPEDEPVLVLA 262
Db 476 LGGPMRSNRSHMSISQRIEANAFAERVPCTGPIPLQMLSSDGMICEDGDFVMA 535
OY 263 LHSLSASEMFAAPGRCAAGMHLFYVRAGYSRMLRLFLAMEKGRHMEY-ECPPYLVY 321
Db 536 AVTTHSSDVFAPPSRLDDGLIYVIRGVSRHQLNFMNLNAGTHLPAGEDPFTKV 595
OY 322 VVVAFRLEPKDKGVFAVDGELVSEAVOGVHP 356
Db 596 VPCRAFRIEPPSSDDILVVDGERVEXPIQAEVMP 630

```

```

RESULT 13
O87OC1 PRELIMINARY: PRT: 661 AA.
AC O87OC1:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE LD11247P.
GN CG2159.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacheb J., Paragas V., Park S., Phouenavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Cejner S.
RA Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
RL EMBL: AY069417; AAL3562.1;
DR SEQUENCE 661 AA: 73870 MW: 426E191EEF82D08 CRC64;
SQ

```

```

Query Match 24.5%: Score 494; DB 5; Length 661;
Best Local Similarity 29.7%: Pred. No. 4,5e-35;
Matches 131; Conservative 62; Mismatches 144; Indels 104; Gaps 8;

```

```

OY 16 RVLVLLNBRGKGKALQDFRSHVQPLAEISFTLMTLERRNHAARELVRSSEELGRMAL 75
Db 216 RVLVLLNBRGSGDAREFVNMHVPLVNEAEVPLDYLT KSNFAIEFLSTFCIDAMCCV 275
OY 76 VMSGSDGLMHEVNNLMRPPMERTAI-QKPLCSLPAGSGNALAALNHYAGTEVYTNEDL 134
Db 276 VAVGSDGLMHEVNNLMRPPMERTAI-QKPLCSLPAGSGNALAALNHYAGTEVYTNEDL 134
OY 135 LTNCTLLCRPVLSFPMNLISLHTASGLRS-FVSLSLANGFTADVLESQKRYRLGEMRFT 193
Db 336 GAALVVISGRS--SPMDVVRVQLQS--RSLSFSLSIGNGILSDVIDESERIMLGIOKFT 391
OY 134 LCTFLRLAALTKYRGRLATLPVGRVGRKTPASVYVVO-----GPVD- 235
Db 392 VWTFLRYLVNLTATYNGRISYLLTLDHEVSTHSATGAAQRMQSSRSCWTHIDMLNGPAPI 451
OY 236 -----AHLVPLEEDEVPSHMO----- 250
Db 452 YHSSAEVLPQEFADVISLETSTINOSFRSRCDSWLSGGSRSEFYYSISESIYHLSADESEF 511

```

```

OY 251 -----VPPEDFVLVLLHSLASEMFAAP 276
Db 512 AGLAASLENRQONTGPASELPDLNPEPSDEQGLVBERGVMAHVAQTHLGIDCHRAP 571
OY 277 MGRCAAGVMHLYFYVRAGYSRAMLRLFLAMEKGRHM-EYECPPYLVYVVAFRLEPKDKG 335
Db 572 KAQLNDGTHLYLLINAGISFRLNMSGTHLPESHDDHVVLPVRAFRLEPYDNH 631
OY 336 GVFAVDGELVSEAVOGVHP 356
Db 632 GIIVVDGERVEXPIQAEVLP 652

```

RESULT 14

```

O9VZW0 PRELIMINARY: PRT: 907 AA.
AC O9VZW0:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE CG2159 protein.
GN CG2159.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Goeayne J.D.,
RA Amatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Bendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoskins R.A., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimbel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasero P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Paie J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers R.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003476; AAF47706.1;
DR Flybase: FBgn0035391; CG2159.
DR Interpro: IPR001251; CRAL_TRIO.

```


[illegible]

acquired immunodeficiency disease; AIDS; rheumatoid arthritis; microbial infection; trauma.

Db 61 RELVSEELGRMDALVVMGSGGLMEHVNVNGIMEREEDMETIQKPLCSLPRGSGNALPAASL 12

OY 121 NHYAGYEQVNTNEDLLTNCITLLCRPVLSLHTRASGLRSFVSLANGFTADVDLE 180
CC |||||||
CC 121 NHYAGYEQVNTNEDLLTNCITLLCRPVLSLHTRASGLRSFVSLANGFTADVDLE 180
CC |||||||
OY 181 SDKYRRLEGEMFTLGTFLRLAALRTYRGRATLTPVGRVGFPTASPVPVVOGPPVDAHLVP 240
CC |||||||
CC 181 SEKRYRLEGEMFTLGTFLRLAALRTYRGRATLTPVGRVGFPTASPVPVVOGPPVDAHLVP 240
CC |||||||
OY 241 LEEQVPSHMVQVPPEDFVLYLALHSHLASMFAPMGRCAGVNHLEFYVAGVSRAML 300
CC |||||||
CC 241 LEEQVPSHMVQVPPEDFVLYLALHSHLASMFAPMGRCAGVNHLEFYVAGVSRAML 300
CC |||||||
OY 301 RLFLAMEGRHMEYECPLVLYVVPVAFRLPEPKDGKGVAVNGELMVSEAVOGVHPNPF 360
CC |||||||
CC 301 RLFLAMEGRHMEYECPLVLYVVPVAFRLPEPKDGKGVAVNGELMVSEAVOGVHPNPF 360
CC |||||||
OY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
CC |||||||
CC 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
CC |||||||

RESULT 3
AAB93955
ID AAB93955 standard; Protein: 384 AA.

AC AAB93955;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:13997.

DE Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PD 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PS Claim 8; SEQ ID 13997; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

CC Sequence 384 AA;

CC Query Match 97.4%; Score 1962; DB 22; Length 384;
CC Best Local Similarity 97.7%; Pred. No. 9.6e-205;
CC Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 MDPAGGPRGVLPKRCRVLYLNLNPRGKGKALQLFRSHVQPLAEAEISFTLTERRNA 60
CC |||||||

DB 1 MDPAGGPRGVLPKRCRVLYLNLNPRGKGKALQLFRSHVQPLAEAEISFTLTERRNA 60
CC |||||||

OY 61 RELVSEELGRMDALVYNSGDGLMHEVYNGLMERPDWETA1QKPLCSLPAGSGNALASL 120
CC |||||||

DB 61 RELVSEELGRMDALVYNSGDGLMHEVYNGLMERPDWETA1QKPLCSLPAGSGNALASL 120
CC |||||||

OY 121 NHYAGYEQVNTNEDLLTNCITLLCRPVLSLHTRASGLRSFVSLANGFTADVDLE 180
CC |||||||

DB 121 NHYAGYEQVNTNEDLLTNCITLLCRPVLSLHTRASGLRSFVSLANGFTADVDLE 180
CC |||||||

OY 181 SDKYRRLEGEMFTLGTFLRLAALRTYRGRATLTPVGRVGFPTASPVPVVOGPPVDAHLVP 240
CC |||||||

DB 181 SDKYRRLEGEMFTLGTFLRLAALRTYRGRATLTPVGRVGFPTASPVPVVOGPPVDAHLVP 240
CC |||||||

OY 181 SEKRYRLEGEMFTLGTFLRLAALRTYRGRATLTPVGRVGFPTASPVPVVOGPPVDAHLVP 240
CC |||||||

DB 181 SEKRYRLEGEMFTLGTFLRLAALRTYRGRATLTPVGRVGFPTASPVPVVOGPPVDAHLVP 240
CC |||||||

OY 241 LEEQVPSHMVQVPPEDFVLYLALHSHLASMFAPMGRCAGVNHLEFYVAGVSRAML 300
CC |||||||

DB 241 LEEQVPSHMVQVPPEDFVLYLALHSHLASMFAPMGRCAGVNHLEFYVAGVSRAML 300
CC |||||||

OY 301 RLFLAMEGRHMEYECPLVLYVVPVAFRLPEPKDGKGVAVNGELMVSEAVOGVHPNPF 360
CC |||||||

DB 301 RLFLAMEGRHMEYECPLVLYVVPVAFRLPEPKDGKGVAVNGELMVSEAVOGVHPNPF 360
CC |||||||

OY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
CC |||||||

DB 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
CC |||||||

RESULT 4
AAB94589

ID AAB94589 standard; Protein: 384 AA.

AC AAB94589;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:15394.

DE Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PD 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INSTR.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length CDNA's defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length CDNA's -

XX Claim 8; SEQ ID 15394; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length CDNA's defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length CDNA's. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length CDNA's. The primers allow obtaining of the full-length
CC CDNA's easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human CDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 384 AA;

XX Query Match 97.4%; Score 1962; DB 22; Length 384;

XX Best Local Similarity 97.7%; Pred. No. 9,6e-205; Indels 0; Gaps 0;

XX Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

XX 1 MDPAGRGVLPKPCRVVLVILNPRGKGKALQIFRSHVOPPLAAEISFTLMTERRNHA 60
DB 1 MDPAGRGVLPKPCRVVLVILNPRGKGKALQIFRSHVOPPLAAEISFTLMTERRNHA 60
XX
XX 61 RELVRSSEELGRMDALVVMGSDGIMHEVNGIMERPDMETAIQPLCSLPAGSGNALASL 120
DB 61 RELVRSSEELGRMDALVVMGSDGIMHEVNGIMERPDMETAIQPLCSLPAGSGNALASL 120
XX
XX 121 NHYAGTQVYNEDILNCTLLICRPLSPMNLSLHTASGSRFSVLSLAFIADVDLE 180
DB 121 NHYAGTQVYNEDILNCTLLICRPLSPMNLSLHTASGSRFSVLSLAFIADVDLE 180
XX
XX 121 NHYAGTQVYNEDILNCTLLICRPLSPMNLSLHTASGSRFSVLSLAFIADVDLE 180
DB 121 NHYAGTQVYNEDILNCTLLICRPLSPMNLSLHTASGSRFSVLSLAFIADVDLE 180
XX
XX 181 SDYKRLGEMRFTLGTFLRLAALTYRGLATLPVGRVGFKTPASPVVVOGPDVAHLYP 240
DB 181 SDYKRLGEMRFTLGTFLRLAALTYRGLATLPVGRVGFKTPASPVVVOGPDVAHLYP 240
XX
XX 241 LEEOVPSHQVYVDEDFVYLALHSHLSAEMPAAGMGCAAGVHMLFVYRAGVSRAML 300
DB 241 LEEOVPSHQVYVDEDFVYLALHSHLSAEMPAAGMGCAAGVHMLFVYRAGVSRAML 300
XX
XX 241 LEEOVPSHQVYVDEDFVYLALHSHLSAEMPAAGMGCAAGVHMLFVYRAGVSRAML 300
DB 241 LEEOVPSHQVYVDEDFVYLALHSHLSAEMPAAGMGCAAGVHMLFVYRAGVSRAML 300
XX
XX 301 RLFLAMEGRHMEYECFYLYVPEVVAFLRLEPKDKGVFAVDGELMVSEAVOGVHNFTW 360
DB 301 RLFLAMEGRHMEYECFYLYVPEVVAFLRLEPKDKGVFAVDGELMVSEAVOGVHNFTW 360
XX
XX 361 MVSGCVPEPPSPKQMPPEEPL 384
DB 361 MVSGCVPEPPSPKQMPPEEPL 384

RESULT 5

AAE00924
ID AAE00924 standard; Protein: 384 AA.

XX AAE00924;

XX 04-JUL-2001 (first entry)

XX Human sphingosine kinase type I (hskl).

XX Human: sphingosine kinase type I; sk1; chromosome 17q25.2;
XX sphingosine-1-phosphate; SIP; drug screening; therapy; haemostasis;
XX thrombosis; allergic reaction; proliferative disease; cancer;
XX haematopoietic disorder; leukaemia; cardiovascular disease; stroke;
XX atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
XX autoimmune disease; inflammatory disease; multiple sclerosis;
XX T helper-1 related disease; chronic obstructive pulmonary disease;
XX asthma; myocardial infarction; neurodegenerative disorder;
XX wound healing; embryogenesis; anticoagulant; cerebroprotective;
XX neuroprotective; antiproliferative; antitumour; cytostatic; cardiant;
XX vulnerability.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Region 16..95 /Label=Conserved_region

XX MO200131029-A2.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-EP09498.

XX 28-OCT-1999; 99US-0162307.

XX 07-FEB-2000; 2000US-0180525.

XX (WARN) WARNER LAMBERT CO.

XX Allen J, Gosink M, Melendez AJ, Takacs L;

XX WPI; 2001-300510/31.

XX N-PSDB; AAD04477.

XX New human sphingosine kinase type I gene for screening drug candidates
XX particularly inhibitors used for preventing or treating e.g.
XX atherosclerosis, thrombosis, asthma and diabetes

XX Claim 3; Fig 1; 91pp; English.

XX The present sequence is human sphingosine kinase type I (hskl).
XX The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the
XX substrate sphingosine to sphingosine-1-phosphate (SIP). The sk1 gene
XX and encoded polypeptide are applicable in screening drug candidates
XX particularly inhibitors for preventing or treating disorders such as
XX haemostasis, thrombosis, allergic reactions, proliferative diseases
XX including cancer, haematopoietic disorders such as leukaemia,
XX cardiovascular diseases such as stroke, atherosclerosis and coronary
XX artery disease, dyslipidaemia, diabetes including type I and type II
XX diabetes, autoimmune and inflammatory diseases such as multiple
XX sclerosis, asthma, myocardial infarction, neurodegenerative disorders,
XX natural wound healing processes and embryogenesis.

XX Sequence 384 AA;

XX Query Match 97.4%; Score 1962; DB 22; Length 384;

XX Best Local Similarity 97.7%; Pred. No. 9,6e-205; Indels 0; Gaps 0;

XX Matches 375; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

XX 1 MDPAGRGVLPKPCRVVLVILNPRGKGKALQIFRSHVOPPLAAEISFTLMTERRNHA 60
DB 1 MDPAGRGVLPKPCRVVLVILNPRGKGKALQIFRSHVOPPLAAEISFTLMTERRNHA 60

QY 61 RELVSEELGRMDALVYMSGDLMEHYVNGIMERPDMETAIOKPLCSLPAGSGNALAASL 120
 DB 61 RELVSEELGRMDALVYMSGDLMEHYVNGIMERPDMETAIOKPLCSLPAGSGNALAASL 120
 QY 121 NHYAGYEQVYVEDLTNCTLLCRVLSPMNLSLHTASGLRFSVSLANGFIADVDLE 180
 DB 121 NHYAGYEQVYVEDLTNCTLLCRVLSPMNLSLHTASGLRFSVSLANGFIADVDLE 180
 QY 181 SDKYRRLGEMRFTLTGFLRLAALRTYRGRLATLTVGRGFKTPASPVVVOOGPVDAHLVP 240
 DB 181 SEKYRRLGEMRFTLTGFLRLAALRTYRGRLATLTVGRGFKTPASPVVVOOGPVDAHLVP 240
 QY 241 LEEQVPSHMVYVPDEDFVLVLLSHLSLASEMFAAPMGRCAGVMHLFFYVAGVSRAML 300
 DB 241 LEEVPSHMVYVPDEDFVLVLLSHLSLASEMFAAPMGRCAGVMHLFFYVAGVSRAML 300
 QY 301 RLFLAMEGRHMEYECPLVYVYVAFRLBPKDKGVFAVDGELMVSEAVOGVHPNYFW 360
 DB 301 RLFLAMEGRHMEYECPLVYVYVAFRLBPKDKGVFAVDGELMVSEAVOGVHPNYFW 360
 QY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
 DB 361 MVSGCVPPEPPSMKPOQMPPEEPL 384

RESULT 6
 AAB48007
 ID AAB48007 standard; Protein: 384 AA.

AC AAB48007;
 DT 19-MAR-2001 (first entry)

XX Human sphingosine kinase protein sequence.

XX Sphingosine kinase; antiarthritic; antisthmatic; antiarteriosclerotic;
 KW antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
 human.

XX Homo sapiens.

XX WO200070028-A1.

XX 23-NOV-2000.

XX 12-MAY-2000; 2000WO-AU00457.

XX 13-MAY-1999; 99AU-0000339.

XX 08-JUL-1999; 99AU-0001504.

XX (JOHN J) JOHNSON & JOHNSON RES PTY LTD.

XX Pltson SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vadas MA;

XX WPI; 2001-016227/02.

XX N-PSDB; AAC84161.

XX Novel sphingosine kinase protein and nucleic acid molecules for
 PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,
 PT atherosclerosis, inflammation, meningitis, multiple sclerosis and
 PT septic shock

XX Claim 9; Fig 7a; 100pp; English.

XX This represents a human sphingosine kinase (SK) protein. The human SK
 CC protein, encoding nucleic acids and modulators are useful for modulating
 CC expression, functional activity or cellular functional activity of
 CC sphingosine kinase in a subject and also for treating a mammal by
 CC modulating the activity of SK. Diseases treated by regulating SK
 CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
 CC inflammation, meningitis, multiple sclerosis and septic shock.

XX Sequence 384 AA;

Query Match 97.2%; Score 1959; DB 22; Length 384;
 Best Local Similarity 97.4%; Pred. No. 2e-204;
 Matches 374; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MDPAGGPRGYLPRRCRYLVLLNPRGKGKALQLEFRSHVQPLLAEBEISFTLMTERRNHA 60
 DB 1 MDPAGGPRGYLPRRCRYLVLLNPRGKGKALQLEFRSHVQPLLAEBEISFTLMTERRNHA 60
 QY 61 RELVSEELGRMDALVYMSGDLMEHYVNGIMERPDMETAIOKPLCSLPAGSGNALAASL 120
 DB 61 RELVSEELGRMDALVYMSGDLMEHYVNGIMERPDMETAIOKPLCSLPAGSGNALAASL 120
 QY 121 NHYAGYEQVYVEDLTNCTLLCRVLSPMNLSLHTASGLRFSVSLANGFIADVDLE 180
 DB 121 NHYAGYEQVYVEDLTNCTLLCRVLSPMNLSLHTASGLRFSVSLANGFIADVDLE 180
 QY 181 SDKYRRLGEMRFTLTGFLRLAALRTYRGRLATLTVGRGFKTPASPVVVOOGPVDAHLVP 240
 DB 181 SEKYRRLGEMRFTLTGFLRLAALRTYRGRLATLTVGRGFKTPASPVVVOOGPVDAHLVP 240
 QY 241 LEEQVPSHMVYVPDEDFVLVLLSHLSLASEMFAAPMGRCAGVMHLFFYVAGVSRAML 300
 DB 241 LEEVPSHMVYVPDEDFVLVLLSHLSLASEMFAAPMGRCAGVMHLFFYVAGVSRAML 300
 QY 301 RLFLAMEGRHMEYECPLVYVYVAFRLBPKDKGVFAVDGELMVSEAVOGVHPNYFW 360
 DB 301 RLFLAMEGRHMEYECPLVYVYVAFRLBPKDKGVFAVDGELMVSEAVOGVHPNYFW 360
 QY 361 MVSGCVPPEPPSMKPOQMPPEEPL 384
 DB 361 MVSGCVPPEPPSMKPOQMPPEEPL 384

RESULT 7
 AAM40180
 ID AAM40180 standard; Protein: 384 AA.

XX AAM40180;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3325.

XX Human; nootropic; immunosuppressant; cytosolic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.
 DR N-PSDB: AA159336.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 5; SEQ ID NO 3325; 10078bp; English.
 PS
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA438642-AA442213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX Sequence 384 AA:
 SQ
 Query Match 97.1%; Score 1957; DB 22; Length 384;
 Best Local Similarity 97.4%; Pred. No. 3.4e-204;
 Matches 374; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MDPAGPRGVLPKPCRVLYLNLNPRGKGRKALQFRSHVQPLLAEEISFTMLTERRNA 60
 DB 1 MDPAGPRGVLPKPCRVLYLNLNPRGKGRKALQFRSHVQPLLAEEISFTMLTERRNA 60
 QY 61 RELVNSEELGRMDALVYMSGDGLMHEVNGLMERPDMETAIQKPLCSPAGSGNALAASL 120
 DB 61 RELVNSEELGRMDALVYMSGDGLMHEVNGLMERPDMETAIQKPLCSPAGSGNALAASL 120
 QY 121 NHVAGYEQVTNEDLTNCTLLCRPYLSPMNLISLHTASGLRSFSVLSLWGFADVDLE 180
 DB 121 NHVAGYEQVTNEDLTNCTLLCRPYLSPMNLISLHTASGLRSFSVLSLWGFADVDLE 180
 QY 181 SDKYRRLGEMKFTLGTFLRLAALRTYRGRLATLPVGRGFKTPASPVVVOGQPVDAHLVP 240
 DB 181 SEKYYRRLGEMKFTLGTFLRLAALRTYRGRLATLPVGRGSKTIPASPVVVOGQPVDAHLVP 240
 QY 241 LEEQVPSHMVQVPEDEFLVLYLALHSHSLASEMFAAPMGRCAAGVNHLEFYVRAGVSRAML 300
 DB 241 LEEQVPSHMVQVPEDEFLVLYLALHSHSLASEMFAAPMGRCAAGVNHLEFYVRAGVSRAML 300
 QY 301 RLFLAMEKGRHMEYECPLYVYVVAFRLEPKDGKGFVAVDGLMAYSEAVOGQVHNNTFM 360
 DB 301 RLFLAMEKGRHMEYECPLYVYVVAFRLEPKDGKGFVAVDGLMAYSEAVOGQVHNNTFM 360
 QY 361 MVSQGVPEPPPSMKPQOMPPPEEPL 384
 DB 361 MVSQGVPEPPPSMKPQOMPPPEEPL 384
 RESULT 8
 ID ABB08089 standard; protein: 384 AA.
 XX
 AC ABB08089;
 XX
 DT 10-SEP-2002 (first entry)
 XX
 DE Human sphingosine kinase 1 (hSPHK1) protein sequence.
 XX
 KW Sphingosine kinase; SPHK; SPHK1; cytostatic; vasotropic; antidiabetic;
 KM neuroprotective; human; enzyme.

XX Homo sapiens.
 OS
 XX US2002042358-A1.
 PN
 XX 11-APR-2002.
 PD
 XX 02-MAR-2001; 2001US-0796487.
 PF
 XX 02-MAR-2000; 2000US-186352P.
 PR
 XX (SPIE/) SPIEGEL S.
 PA
 XX Spiegel S;
 PI
 XX WPI: 2002-478846/51.
 DR
 XX New isolated sphingosine kinase, useful in identifying modulators for
 PT treating e.g. cancer, also related nucleic acid, vectors and
 PT transformed cells -
 PT disclosure: Fig 3; 24pp; English.
 PS
 XX The invention relates to an isolated sphingosine kinase (SPHK) DNA. Cells
 CC transformed with SPHK DNA are used to screen for agents that reduce,
 CC eliminate or promote SPHK activity. Agents that inhibit activity are
 CC useful for decreasing cell proliferation, e.g. for treating cancer, and
 CC for treating diseases associated with abnormal migration and motility of
 CC cells, e.g. restenosis or diabetic neuropathy. Agents that increase
 CC activity are used to reduce cell death. Antibodies raised against SPHK,
 CC and primers or oligonucleotides derived from the DNA are useful for
 CC diagnosis. The antibodies are also useful as therapeutic inhibitors. The
 CC present sequence represents a human sphingosine kinase 1 (hSPHK1).
 CC
 XX Sequence 384 AA:
 SQ
 Query Match 97.0%; Score 1954; DB 23; Length 384;
 Best Local Similarity 96.9%; Pred. No. 7.2e-204;
 Matches 372; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MDPAGPRGVLPKPCRVLYLNLNPRGKGRKALQFRSHVQPLLAEEISFTMLTERRNA 60
 DB 1 MDPAGPRGVLPKPCRVLYLNLNPRGKGRKALQFRSHVQPLLAEEISFTMLTERRNA 60
 QY 61 RELVNSEELGRMDALVYMSGDGLMHEVNGLMERPDMETAIQKPLCSPAGSGNALAASL 120
 DB 61 RELVNSEELGRMDALVYMSGDGLMHEVNGLMERPDMETAIQKPLCSPAGSGNALAASL 120
 QY 121 NHVAGYEQVTNEDLTNCTLLCRPYLSPMNLISLHTASGLRSFSVLSLWGFADVDLE 180
 DB 121 NHVAGYEQVTNEDLTNCTLLCRPYLSPMNLISLHTASGLRSFSVLSLWGFADVDLE 180
 QY 181 SDKYRRLGEMKFTLGTFLRLAALRTYRGRLATLPVGRGFKTPASPVVVOGQPVDAHLVP 240
 DB 181 SEKYYRRLGEMKFTLGTFLRLAALRTYRGRLATLPVGRGSKTIPASPVVVOGQPVDAHLVP 240
 QY 241 LEEQVPSHMVQVPEDEFLVLYLALHSHSLASEMFAAPMGRCAAGVNHLEFYVRAGVSRAML 300
 DB 241 LEEQVPSHMVQVPEDEFLVLYLALHSHSLASEMFAAPMGRCAAGVNHLEFYVRAGVSRAML 300
 QY 301 RLFLAMEKGRHMEYECPLYVYVVAFRLEPKDGKGFVAVDGLMAYSEAVOGQVHNNTFM 360
 DB 301 RLFLAMEKGRHMEYECPLYVYVVAFRLEPKDGKGFVAVDGLMAYSEAVOGQVHNNTFM 360
 QY 361 MVSQGVPEPPPSMKPQOMPPPEEPL 384
 DB 361 MVSQGVPEPPPSMKPQOMPPPEEPL 384
 RESULT 9
 ID AA196057 standard; protein: 384 AA.
 XX
 AA196057
 ID AA196057 standard; protein: 384 AA.
 XX

XX	AAV96057;
XX	
DT	05-DEC-2000 (first entry)
XX	
DE	Human sphingosine kinase A.
XX	
KW	Sphingosine kinase A; SKA; human; drug screening; infection;
KW	antiinflammatory; antiallergic; anticancer; inflammation; allergy;
KW	cancer; therapy; diagnosis.
XX	
OS	Homo sapiens.
PN	WO200052173-A2.
XX	
PD	08-SEP-2000.
XX	
PF	02-MAR-2000; 2000MO-CA00223.
XX	
PR	02-MAR-1999; 99US-0122516.
XX	
PA	(ALLX) NPS ALLELIX CORP.
XX	
PI	Munroe D, Gupta A, Falzone GR;
XX	
DR	WPI; 2000-572185/53.
DR	N-PSDB; AAA50508.
XX	
PT	New human sphingosine kinase A, B and C polynucleotides and
PT	polypeptides useful in e.g. chromosome and gene mapping, and detecting
PT	inflammation or disease associated with abnormal levels of sphingosine
PT	kinase expression
XX	
XX	Disclosure; Fig 3; 81pp; English.

	CC	The present sequence is that of human sphingosine kinase A (SKA),
	CC	an enzyme that phosphorylates sphingosine to form sphingosine
	CC	1-phosphate. The sequence was deduced from that of a polynucleotide
	CC	(see AA50508) isolated from an HeLa cDNA library. The invention
	CC	provides polynucleotides (see AA50508-10) and polypeptides (see
	CC	AA56057-59) for the human sphingosine kinase (SK) homologues SKA,
	CC	SKB and SKC. The polypeptides can be obtained using recombinant
	CC	DNA methods, and host cells containing expression vectors including
	CC	SK polynucleotides are used in a claimed method of screening for
	CC	compounds that inhibit or activate human SK activity. Human SK
	CC	specific antibodies, inhibitors, ligands or their analogues can be
	CC	used as bioactive agents to treat inflammation or disease including
	CC	viral, bacterial or fungal infections, allergic responses,
	CC	mechanical injury associated with trauma, hereditary diseases,
	CC	Lymphoma or carcinoma, and other conditions with activate the
	CC	genes of kidney, lung, heart, lymphoid or tissues of the nervous
	CC	system.
	XX	
SQ	Sequence	384 AA;
OY	Query Match	96.6%; Score 1946; DB 21; Length 384;
	Best Local Similarity	96.9%; Pred. No. 5.3e-203;
MATCHES	372; Conservative	2; Mismatches 10; Indels 0; Gaps 0;
DB	1	MDPAGGPGGVLPDRCRVYLNLNPRGGKKALQLFRSHVQPLLAKEAFISFTLMTFRRRNA 60
	1	MDPPGGPGGVLPDPCRVYVLNLNPRGGKKAQLDFRSHVQPLLAKEAFISFTLMTERRHNA 60
OY	61	RELVSSELGLKMDALVWMSGDGLMHEVVNGLMERPDWETAIOKPLCSIPAGSGNALASL 120
DB	61	RELVRSSELGLKMDALVWMSGDGLMHEVVNGLMERPDWETAIOKPLCLIPAGSGNALASL 120
OY	121	NHYAVGEQVTNEDLLTNCTLLLCGRPVLSPMNILSLHTASGRLSFVSLSIANGFTADVLE 180
DB	121	NHYAVGEQVTNEDLLTNCTLLLCGRRLSPMNILSLHTASGLRLFVSLSIANGFTADVLE 180
OY	181	SDKRRRLGLEMPFTLGTFRLRALAKTRYKRRLATLPYGRGVEFTPASPVVVOOGPDVAHLVP 240
DB	181	SEKVRRLGLEMPFTLTGTFRLRAALKTRYKRRLAYLTPVGKVGSTFPASPVVVOOGPDVAHLVP 240

QY	24	LEEOGVSHQVYDEDFEVLVLLSHSLASEMFAAPMQRCAAGVHLEFYVAGVSRAML	3000
Db	241	LEEPVSHMTVPDEDFEVLVLLSHSLASEMFAAPMQRCAAGVHLEFYVAGVSRAML	3000
QY	301	RLEFLAEKGGHMECECPFLVYPPVVAFLPEKDGKGVAVDGLMVAEAVOGQVHPYFW	3600
Db	301	RLEFLAEKGGHMECECPFLVYPPVVAFLPEKDGKGVAVDGLMVAEAVOGQVHPYFW	3600
QY	361	MVSGCYEPPPSMKPQOMPPEEPL	384
Db	361	MVSGCYEPPPSMKPQOMPPEEPL	384
RESULT 10			
ABB90209	ID	ABB90209 standard; Protein; 333 AA.	
XX	AC	ABB90209;	
XX	DE	24-MAY-2002 (first entry)	
XX	DE	Human polypeptide SEQ ID NO 2585.	
KW	XX	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;	
KW	XX	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;	
KW	XX	vulnerable; anticoagulant; antibacterial; antifungal; antiparasitic;	
KW	XX	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;	
KW	XX	neurological disease; infection; human; secreted protein.	
XX	OS	Homo sapiens.	
PN	XX	WO200190304-A2.	
PD	XX	29-NOV-2001.	
XX	XX	18-MAY-2001; 2001WO-US16450.	
PF	XX	19-MAY-2000; 2000US-205515P.	
PR	XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	XX	Blrse CE, Rosen CA;	
PI	XX	WPI: 2002-122018/16.	
DR	XX	N-PSDB: ABL90618.	
PT	XX	Novel 1405 isolated polypeptides, useful for diagnosis, treatment and	
PT	XX	prevention of neural, immune system, muscular, reproductive,	
PT	XX	gastrointestinal, pulmonary, cardiovascular, renal and proliferative	
PT	XX	disorders -	
XX	XX	Claim 11: SEQ ID NO 2585; 2081bp + Sequence Listing: English.	
XX	XX	The invention relates to novel genes (ABL89449-ABL90853) and proteins	
CC	XX	(ABB89040-ABB90444) useful for preventing, treating or ameliorating	
CC	XX	medical conditions e.g. by protein or gene therapy. The genes are	
CC	XX	isolated from a range of human tissues disclosed in the specification.	
CC	XX	The nucleic acids, proteins, antibodies and (ant)agonists are useful	
CC	XX	in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast	
CC	XX	and ovarian cancer and other cancers of the adrenal gland, bone, bone	
CC	XX	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;	
CC	XX	(b) immune disorders e.g. Addison's disease, allergies, autoimmune	
CC	XX	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's	
CC	XX	disease, multiple sclerosis, rheumatoid arthritis and ulcerative	
CC	XX	colitis; (c) cardiovascular disorders such as myocardial ischaemias;	
CC	XX	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and	
CC	XX	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal	
CC	XX	and parasitic infections.	
CC	XX	Note: The sequence data for this patent did not form part of the	
CC	XX	printed specification, but was obtained in electronic format directly	
CC	XX	from WIPO at ftp.wipo.int/pub/published_pct_sequences .	

SQ Sequence 333 AA:
 Query Match 84.3%; Score 1698; DB 23; Length 333;
 Best Local Similarity 97.3%; Pred. No. 4,6e-176;
 Matches 324; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 52 MTERRHAAELVRSSEIGRWDAIYVSGDGLMHEVYNGIMERPDMETAIQKFLCSIPAG 111
 DB 1 MTERRHAAELVRSSEIGRWDAIYVSGDGLMHEVYNGIMERPDMETAIQKFLCSIPAG 60

QY 112 SGNAALASLNHYAGYEQVETNEDLLTNCITLLCRPVLSPMNLISLHTASGLRSESVLSIAW 171
 DB 61 SGNAALASLNHYAGYEQVETNEDLLTNCITLLCRPVLSPMNLISLHTASGLRSESVLSIAW 120

QY 172 GFADVDVLESDKTRRLGEMFTIGTFLRLAALTYRKRATLVYGRGFTPASPYYVQ 231
 DB 121 GFADVDVLESDKTRRLGEMFTIGTFLRLAALTYRKRATLVYGRGFTPASPYYVQ 180

QY 232 GPVDAHLVPLEEQVPSHMVYVPPDEDFVLVTLALSHSLASEMFAAPMGRCAAGVMHLFYVR 291
 DB 181 GPVDAHLVPLEEQVPSHMVYVPPDEDFVLVTLALSHSLASEMFAAPMGRCAAGVMHLFYVR 240

QY 292 AGVSRAMLRLFLAMEGRMEYECPLYVYVPAVAFRLBPKDKGVFAVDGELMVSBAVQ 351
 DB 241 AGVSRAMLRLFLAMEGRMEYECPLYVYVPAVAFRLBPKDKGVFAVDGELMVSBAVQ 300

QY 352 GQVHPNPFMMVSGCVPEPPSMKPOQMPPEEP 384
 DB 301 GQVHPNPFMMVSGCVPEPPSMKPOQMPPEEP 333

RESULT 11
 AAE07883
 ID AAE07883 standard; Protein: 382 AA.
 AC AAE07883;
 DT 01-NOV-2001 (first entry)
 DE Mouse consensus sphingosine kinase (Sphk) protein.
 KW Mouse; sphingosine kinase; Sphk; restenosis; ischaemia; gene therapy; antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis; cytosolic; tumor necrosis factor-alpha; TNF; atherosclerosis; lymphoma; leukaemia; vasotropic; cell proliferative disorder; vascular disease.
 OS Mus musculus.
 FN WO200160990-A2.
 PD 23-AUG-2001.
 PF 14-FEB-2001; 2001WO-US04789.
 PR 14-FEB-2000; 2000US-0182360.
 PA (CURA-) CURAGEN CORP.
 PA (GETH) GENENTECH INC.
 PI Rastelli L;
 DR WPI: 2001-514770/56.
 DR N-PSDB; AAD14425.
 PT An isolated sphingosine kinase polypeptide useful for treating a Sphk-associated disorder especially cancer, restenosis or ischemia in a human -
 PS Example 1; Page 92-94; 107pp; English.
 CC The present invention relates to sphingosine kinase (Sphk) polypeptides and nucleic acids encoding them. Sphk is useful for treating a Sphk-

CC associated disorder especially cancers such as leukaemia, lymphoma, CC ovarian, breast, lung, colon, testicular, stomach and skin, CC atherosclerosis, restenosis or ischaemia and cell proliferative disease CC or disorder associated with vascular diseases. Sphk gene is used in gene CC therapy and antisense-therapy. Sphingolipids serving as signalling CC molecules, have recently emerged as regulators of cell growth, CC differentiation, diverse cell phenotypes and cell death. Activation of CC Sphk by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human CC endothelial cells. The present sequence is consensus mouse sphingosine CC kinase (Sphk) protein.
 CC
 SQ Sequence 382 AA:
 Query Match 79.1%; Score 1594.5; DB 22; Length 382;
 Best Local Similarity 79.1%; Pred. No. 1.1e-164;
 Matches 303; Conservative 35; Mismatches 44; Indels 1; Gaps 1;

QY 1 MDPAGPBGVLPBPCRYLVILNPRGKGRALDLSRSHVOPPLAEISFTLMTERRHA 60
 DB 1 MEVVECPBGILLPRPCRYLVILNPOGKGRALDLSRSHVOPPLAEISFTLMTERRHA 60

QY 61 RELVSEELGRMDALYVSGDGLMHEVYNGIMERPDMETAIQKFLCSIPAGSNAALASL 120
 DB 61 RELVCAEEELGRMDALYVSGDGLMHEVYNGIMERPDMETAIQKFLCSIPAGSNAALASL 120

QY 121 NHYAGYEQVETNEDLLTNCITLLCRPVLSPMNLISLHTASGLRSESVLSIAWGFIAVDLE 180
 DB 121 NHYAGYEQVETNEDLLTNCITLLCRPVLSPMNLISLHTASGLRSESVLSIAWGFIAVDLE 180

QY 181 SDKYRRLGEMFTIGTFLRLAALTYRKRATLVYGRGFTPASPYYVQGVDAHLVY 240
 DB 181 SEKYRRLGEMFTIGTFLRLAALTYRKRATLVYGRGFTPASPYYVQGVDAHLVY 239

QY 241 LEEQVPSHMVYVPPDEDFVLVTLALSHSLASEMFAAPMGRCAAGVMHLFYVRAVSRAILL 300
 DB 240 LEEVPSHMTVVEQDPLVTLVTLHLSSELPAAPMGREAGVMHLFYVRAVSRAILL 299

QY 301 RLFLAMEGRMEYECPLYVYVPAVAFRLBPKDKGVFAVDGELMVSBAVQGVHPNYFM 360
 DB 300 RLFLAMGKGMHMLDCPYLVHPVVAFRLEPRSGRGVSDGLMCEAVQGVHPNYFM 359

QY 361 MVSGCVPEPPSMKPOQMPPEEP 383
 DB 360 MVGSRDAPGRDSRRGPPEEP 382

RESULT 12
 AAY56054
 ID AAY56054 standard; Protein: 388 AA.
 AC AAY56054;
 DT 28-MAR-2000 (first entry)
 DE Mouse sphingosine kinase 1b protein.
 KW Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis; antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis; cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis.
 OS Mus musculus.
 FN WO9961581-A2.
 PD 02-DEC-1999.
 PF 25-MAY-1999; 99WO-US11521.
 PR 26-MAY-1998; 98US-0086657.
 PR 11-AUG-1998; 98US-0096049.
 PA (DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.

RESULT 14
ID AAY56053 standard; protein; 381 AA.
XX AAY56053;
AC AAY56053;
XX 28-MAR-2000 (first entry)
XX Mouse sphingosine kinase 1a protein.
DE
XX Anticancer; antiproliferative; antiatherosclerotic; stroke; restenosis;
KM antineurodegeneration; sphingosine kinase; cell proliferation; apoptosis;
KM cancer; diabetic neuropathy; Alzheimer's disease; atherosclerosis.
XX
OS Mus musculus.
XX W0961581-A2.
XX W0961581-A2.
XX 02-DEC-1999.
XX 25-MAY-1999; 99MO-US11521.
XX 26-MAY-1998; 98US-0086657.
XX 11-AUG-1998; 98US-0096049.
XX (DEAN-) OFFICE DEAN RES & GRADUATE EDUCATION.
XX Spiegel S;
XX WPI; 2000-072612/06.
XX N-PSDB; AAZ47167.
XX New sphingosine kinase, used to treat diseases involving abnormal cell
PT proliferation, e.g. cancer.
XX
XX Example 1; Fig 1; 116pp; English.
XX
XX This sequence represents the mouse sphingosine kinase 1a. The coding
CC sequence was isolated by searching a dbEST (expressed sequence tag
CC database) using amino acid sequence results from sequenced tryptic
CC digests of the rat sphingosine kinase. Expression of sphingosine kinase
CC in cells results in formation of sphingosine-1-phosphate, a known second
CC messenger, and confers serum-independent growth; increases proliferation,
CC and suppresses serum-deprivation or ceramide-induced apoptosis. The
CC sphingosine kinase nucleic acid is used: (a) to increase sphingosine
CC kinase content of cells, specifically for reducing cell death and/or
CC increasing cell proliferation; and (b) to produce transfected cells that
CC are used to screen for agents that inhibit or promote sphingosine kinase
CC activity. Agents that reduce sphingosine kinase activity or expression
CC are used: (i) to reduce cell proliferation, specifically for treating
CC cancer, and (ii) to treat diseases associated with abnormal cell
CC migration or motility, particularly cancer, restenosis or diabetic
CC neuropathy (but also atherosclerosis, stroke and Alzheimer's disease),
CC whereas agents that stimulate sphingosine kinase can be used to treat
CC conditions associated with reduced cell proliferation, e.g. developmental
CC retardation.
XX
XX Sequence 381 AA:
SO
Query Match 78.5%; Score 1582.5; DB 21; Length 381;
Best Local Similarity 79.6%; Pred. No. 2.2e-163;
Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;
QY 7 PRGVLRPPCRVYVLLNPRGKRAQLFRSHVQPLAEATISTMTLTERNNARELYRS 66
Db 6 PRGLPPRCRVYVLLNPRGKRAQLFRSHVQPLAEATISTMTLTERNNARELYRS 65
QY 67 EEELGRMDALVWMSGDGLMHEVYVNGLMERPDMETAIQPLCSLPAGSGNALAASVNHAGY 126
Db 66 EEELGHMDALVWMSGDGLMHEVYVNGLMERPDMETAIQPLCSLPAGSGNALAASVNHAGY 125

QY 127 EGYTNEDLLNCTLLICRPVLSPMNLSLHTASGLRSFVSLAMGFADVDLESDKRR 186
Db 126 EGYTNEDLLNCTLLICRRRLSPMNLSTHTASGLRLYSLSMGFADVDESEKRR 185
QY 187 LGEHRTLTGFLALARTYRGLATLPVGRNGCFKTPASPVVVOGPDVAHLPLEEYVP 246
Db 186 LGEIRFTVGTFFFLASRLTYOGOLAVLPVGTASKRPAS-TLVQKSPVTHLPLEEPVP 244
QY 247 SHMOVVPDEDFVYVLLNLSHSLASEMFAPMGSCAGVHMLFVRAGVSRAMLRLFLAM 306
Db 245 SHMTVPEODEFLVLLVLTHTLSSELPAAMGRCGEVHMLFVRAGVSRAALLRLFLAM 304
QY 307 EKGRRHMEDECPYLVYVVAFRLEPKDGKVFVDELMWSEAVOGVPHNFMVSGCV 366
Db 305 QKQKHHELDCLPLVHVPVAFRLPRRSQGRVFSVDDELWCEAVOGVPHNFMVCGSR 364
QY 367 EPPPSMKPOOMPPEEP 383
Db 365 DAPSGRDSRRGPPEEP 381

RESULT 15
ID ABB08087 standard; protein; 381 AA.
XX ABB08087;
AC ABB08087;
XX 10-SEP-2002 (first entry)
XX
XX Murine sphingosine kinase 1a (mSPHK1a) protein sequence.
DE
XX Sphingosine kinase; SPHK; SPHK1a; cytosolic; vasotropic; antidiabetic;
KM neuroprotective; mouse; enzyme.
XX
XX Mus musculus.
XX US2002042358-A1.
XX 11-APR-2002.
XX 02-MAR-2001; 2001US-0796487.
XX 02-MAR-2000; 2000US-186352P.
XX (SPIE/) SPIEGEL S.
XX Spiegel S;
XX WPI; 2002-478846/51.
XX New isolated sphingosine kinase, useful in identifying modulators for
PT treating e.g. cancer, also related nucleic acid, vectors and
PT transformed cells.
XX
XX Disclosure; Fig 1; 24pp; English.
XX
XX The invention relates to an isolated sphingosine kinase (SPHK) DNA. Cells
CC transformed with SPHK DNA are used to screen for agents that reduce,
CC eliminate or promote SPHK activity. Agents that inhibit activity are
CC useful for decreasing cell proliferation, e.g. for treating cancer, and
CC for treating diseases associated with abnormal migration and motility of
CC cells, e.g. restenosis or diabetic neuropathy. Agents that increase
CC activity are used to reduce cell death. Antibodies raised against SPHK,
CC and primers or oligonucleotides derived from the DNA are useful for
CC diagnosis. The antibodies are also useful as therapeutic inhibitors. The
CC present sequence represents a murine sphingosine kinase 1a (mSPHK1a)
CC corresponding to residues 124 to 504 of SPHK1a in Genbank Accession
CC no. AAC61697.
XX
XX Sequence 381 AA:
SO
Query Match 78.5%; Score 1582.5; DB 23; Length 381;
Best Local Similarity 79.6%; Pred. No. 2.2e-163;

Matches 300; Conservative 35; Mismatches 41; Indels 1; Gaps 1;

```

Oy 7 PRGVLPPRCRYLVNLRGKGAOLFRSHVOPFLAEIETFTLMTERRNHARELYRS 66
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 PRGVLPPRCRYLVNLRGKGAOLFRSHVOPFLAEIETFTLMTERRNHARELYCA 65
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 67 EEIGRMDALVVMGSGDGLMHEVNGLMERPPMETAIQKPLCSLPAGSGNALASLNHYAGY 126
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 EEIGRMDALVVMGSGDGLMHEVNGLMERPPMETAIQKPLCSLPAGSGNALASLNHYAGY 125
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 127 EGYTNEDELINCTILLCRPVLSFPMNLSLHTASGLRSFVSLAMGFIADVDLESDKYRR 186
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 EGYTNEDELINCTILLCRPVLSFPMNLSLHTASGLRSFVSLAMGFIADVDLESDKYRR 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 187 LGEMRFTLGTFLRLAALRTYRGRLATLPVGRVGFKTPASPVVVOGSPVDAHLVPLEEQVP 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 LGEMRFTLGTFLRLAALRTYRGRLATLPVGRVGFKTPASPVVVOGSPVDAHLVPLEEQVP 244
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 247 SHMOVPEDEDFVLVLTALSHSLASEMFAAPMGRCAAGVMHLFYRAGVSRAMLRLFLAM 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 SHMTVPEQDFLVLVLTALSHSLASEMFAAPMGRCAAGVMHLFYRAGVSRAMLRLFLAM 304
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 307 EKGHRMEYECRPLVYVPVVAFRLEPKDKGVFAVDGELMVS EAVQGVHPNYFMVMSGCV 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 OKGKHEMLDCPYLVHVPVVAFRLEPKDKGVFAVDGELMVS EAVQGVHPNYFMVMSGCV 364
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 367 EPPPSMKPQOMPPPEP 383
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 365 DAPSGRDSRRGPPPEP 381
    |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: July 12, 2003, 18:07:30
 Job time : 46 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 16:04:37 ; Search time 273 Seconds

(without alignments)
9248,401 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599

Sequence: 1 cccccggggcctctatagc.....aataaagtgcattcccaa 1599

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1105431 seqs, 789497651 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCIT_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/PCITUS_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1599	100.0	1600	10 US-09-784-810A-1	Sequence 1, Appl1
2	624	39.0	1533	9 US-10-023-282-90	Sequence 90, Appl1
3	433	27.1	1155	10 US-09-970-516-1	Sequence 1, Appl1
4	192	12.0	296	9 US-10-015-219-658	Sequence 658, App
5	192	12.0	296	10 US-09-777-564-658	Sequence 658, App
6	191	11.9	394	10 US-09-954-456-1756	Sequence 1756, Ap
7	175	10.9	199	9 US-09-796-692-2905	Sequence 2905, Ap
8	175	10.9	199	10 US-10-040-862-9248	Sequence 2905, Ap
9	49	3.1	480	10 US-09-783-590-9248	Sequence 9248, Ap
10	30	1.9	30	10 US-09-796-487-17	Sequence 17, Appl
11	25	1.6	1149	10 US-09-970-516-5	Sequence 5, Appl1
12	25	1.6	1759	10 US-09-784-810A-3	Sequence 5, Appl1
13	22	1.4	459	9 US-09-918-995-26954	Sequence 26954, A
14	21	1.3	27	10 US-09-796-487-16	Sequence 16, Appl
15	20	1.3	20	10 US-09-796-487-15	Sequence 15, Appl
16	19	1.2	537	9 US-09-918-995-25081	Sequence 25081, A
17	19	1.2	546	9 US-10-136-761-6084	Sequence 6084, Ap
18	19	1.2	592	10 US-09-864-761-6833	Sequence 6833, Ap
19	19	1.2	1133	9 US-10-282-048-1	Sequence 1, Appl1

20	19	1.2	1515	9 US-10-238-709-1	Sequence 1, Appl1
21	19	1.2	1515	10 US-09-804-471A-1	Sequence 1, Appl1
22	19	1.2	1942	10 US-09-764-877-3173	Sequence 3173, Ap
23	19	1.2	4285	10 US-09-759-622-1	Sequence 1, Appl1
24	19	1.2	5877	12 US-10-028-946-3	Sequence 3, Appl1
25	19	1.2	6159	9 US-10-017-216-3	Sequence 1, Appl1
26	19	1.2	6165	12 US-10-028-946-1	Sequence 1, Appl1
27	19	1.2	6574	9 US-10-017-216-1	Sequence 3, Appl1
28	19	1.2	48763	9 US-10-282-048-3	Sequence 3, Appl1
29	19	1.2	142299	9 US-09-911-077A-14	Sequence 14, Appl
30	19	1.2	174493	9 US-10-238-709-3	Sequence 3, Appl1
31	19	1.2	174493	10 US-09-804-471A-3	Sequence 1, Appl1
32	19	1.2	9025608	9 US-10-156-761-1	Sequence 18007, A
33	18	1.1	153	10 US-09-864-761-18007	Sequence 4219, Ap
34	18	1.1	375	10 US-09-867-701-1219	Sequence 37471, A
35	18	1.1	379	9 US-09-918-995-37471	Sequence 9674, Ap
36	18	1.1	391	10 US-09-783-590-9674	Sequence 37665, A
37	18	1.1	394	9 US-09-918-995-37665	Sequence 26, Appl
38	18	1.1	433	10 US-09-969-347-26	Sequence 7542, Ap
39	18	1.1	435	9 US-09-764-891-7542	Sequence 1526, Ap
40	18	1.1	444	10 US-09-864-761-1526	Sequence 1246, Ap
41	18	1.1	495	10 US-09-864-761-1246	Sequence 6730, Ap
42	18	1.1	517	9 US-09-764-891-6730	Sequence 6732, Ap
43	18	1.1	517	9 US-09-764-891-6732	Sequence 7828, Ap
44	18	1.1	556	10 US-09-864-761-7828	Sequence 8140, Ap
45	18	1.1	585	10 US-09-864-761-8140	

ALIGNMENTS

RESULT 1

US-09-784-810A-1

Sequence 1, Application US/09784810A

Patent No. US20020082203A1

GENERAL INFORMATION:

APPLICANT: BASTELLI, LUCA

TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING

TITLE OF INVENTION: SAME

FILE REFERENCE: 10716-08

CURRENT APPLICATION NUMBER: US/09/784,810A

CURRENT FILING DATE: 2001-02-14

PRIOR APPLICATION NUMBER: 60/182,360

PRIOR FILING DATE: 2000-02-14

PRIOR APPLICATION NUMBER: 60/191,261

PRIOR FILING DATE: 2000-03-22

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 1600

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: modified_base

LOCATION: (1)

OTHER INFORMATION: a, t, c, g, other or unknown

US-09-784-810A-1

Query Match

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1599; Conservative

Mismatches 0; Indels 0; Gaps 0;

QY	1	CCCCCGGGGCTCCATAGCCAGCGCTCCGGGGGGAAGCGGACCCACAGCCGGGCC	60
DB	2	CCCCCGGGGCTCCATAGCCAGCGCTCCGGGGGGAAGCGGACCCACAGCGCGGCC	61
QY	61	TGCGAGCGCCCTTGCGAGCAGCAGCATTAAGAGCTGAAGGAGGAGCGCCCGCCAGGGC	120
DB	62	TGCGAGCGCCCTTGCGAGCAGCAGCATTAAGAGCTGAAGGAGGAGCGCCCGCCAGGGC	121
QY	121	AGCGCCCGCCAGCGCCAGGAGCCCTTGCGAGCGGAGCGCGGGCTGAGGTATGAT	180
DB	122	AGCGCCCGCCAGCGCCAGGAGCCCTTGCGAGCGGAGCGCGGGCTGAGGTATGAT	181

QY	181	CCAGCGGGGGCCCCCGGGGGCGTGTCTCCCGGGCCCTGCGCGTGTGTGTGTAAAC	240
Db	182	CCAGCGGGGGGGCCCCCGGGGGCGTGTCTCCCGGGCCCTGCGCGTGTGTGTGTAAAC	241
QY	241	CCCGGGGGGGCAAGGGCAAGGCGCTTGACACTCTTCCGGAGTACAGCGCAAGCCCTTTTG	300
Db	242	CCCGGGGGGGGGCAAGGGCAAGGCGCTTGACACTCTTCCGGAGTACAGCGCAAGCCCTTTTG	301
QY	301	GGTGAAGGCTGAATTCCTTTCACGCTGATGCTCATTGAGCGGGGAACACAGCGCGGAG	360
Db	302	GGTGAAGGCTGAATTCCTTTCACGCTGATGCTCATTGAGCGGGGAACACAGCGCGGAG	361
QY	361	CTGGGTCCGGGCGGAGAGGCTGGGGCCGCTGGGAGAGGCTGTGGTTCATGTCTGAGACGG	420
Db	362	CTGGGTCCGGGCGGAGAGGCTGGGGCCGCTGGGAGAGGCTGTGGTTCATGTCTGAGACGG	421
QY	421	CTGATCCAGAGGTGTGAACGGGGCTCATGAGACGGGCTGACTGAGGAGACCCGCAATCCAG	480
Db	422	CTGATCCAGAGGTGTGTGAACGGGGCTCATGAGACGGGCTGACTGAGGAGACCCGCAATCCAG	481
QY	481	AAGCCCCCTGTGATAGCCTCCAGAGAGGCTCTGGGAAGCGCGTGGCACTTCTTGAACAT	540
Db	482	AAGCCCCCTGTGATAGCCTCCAGAGAGGCTCTGGGAAGCGCGTGGCACTTCTTGAACAT	541
QY	541	TATGCTGGCTATGAGCAGGTCAACCATGGAAGACCTCTACCAACATGCAAGGATATGAG	600
Db	542	TATGCTGGCTATGAGCAGGTCAACCATGGAAGACCTCTACCAACATGCAAGGATATGAG	601
QY	601	TGCGCGCCGGGTGTGTCAACCCATGAACCTGTGTCTCTGACACAGGGCTTGCGGCGC	660
Db	602	TGCGCGCCGGGTGTGTCAACCCATGAACCTGTGTCTCTGACACAGGGCTTGCGGCGC	661
QY	661	TGCTTCTCTGTCTCAGCTCGGCGCTGGGGCTTCATTGGCATGTGGACCTTAAGATGAT	720
Db	662	TGCTTCTCTGTCTCAGCTCGGCGCTGGGGCTTCATTGGCATGTGGACCTTAAGATGAT	721
QY	721	AAGTATCGGCGTCTGGGGGAGATGCGCTTCACTGTGGCACTTCTCGTGTGGCAGGC	780
Db	722	AAGTATCGGCGTCTGGGGGAGATGCGCTTCACTGTGGCACTTCTCGTGTGGCAGGC	781
QY	781	CTGGGCACTTACCGGGGGCGAGTGGCTACCCCTCTGTAAGAAAGTGGGTTTCAAGACA	840
Db	782	CTGGGCACTTACCGGGGGCGAGTGGCTACCCCTCTGTAAGAAAGTGGGTTTCAAGACA	841
QY	841	CCGTGTCGCCCGTGTGTGTCCAGAGAGGCCCGGTAGATGACACACTGGGTGCACATGGAG	900
Db	842	CCGTGTCGCCCGTGTGTGTCCAGAGAGGCCCGGTAGATGACACACTGGGTGCACATGGAG	901
QY	901	GAGCAGGTGCTTCTTCACTGTGAGGTGTGGTCCCGACGAGAGACTTGTGTCTAGTCTGGA	960
Db	902	GAGCAGGTGCTTCTTCACTGTGAGGTGTGGTCCCGACGAGAGACTTGTGTCTAGTCTGGA	961
QY	961	CTGTGCACTCGCACCTGAGCCAGTAGATGTTTGTGTGACACCATAGGGGCCGCTGTGCACT	1020
Db	962	CTGTGCACTCGCACCTGAGCCAGTAGATGTTTGTGTGACACCATAGGGGCCGCTGTGCACT	1021
QY	1021	GCGCTCATGATCTGTGTTTACGTGGCGGGAGGTGTCTGATGCACTCTGTGTGGCGCTC	1080
Db	1022	GCGCTCATGATCTGTGTTTACGTGGCGGGAGGTGTCTGATGCACTCTGTGTGGCGCTC	1081
QY	1081	TTCTCGGCGCATGAGAAAGGCAAGGCATATGAGATGTAATGCCCTTACTTGGTATATGTG	1140
Db	1082	TTCTCGGCGCATGAGAAAGGCAAGGCATATGAGATGTAATGCCCTTACTTGGTATATGTG	1141
QY	1141	CCCGGTGGCTTCCGTTGGAGGCCAAGATGGGAAGGTGTGTTGCAGTGGATGGG	1200
Db	1142	CCCGGTGGCTTCCGTTGGAGGCCAAGATGGGAAGGTGTGTTGCAGTGGATGGG	1201
QY	1201	GAATTTGATGTATTACGAGGCGGTGCGAGGCCAGGTGCAACCAACTACTTCTGGATGTG	1260
Db	1202	GAATTTGATGTATTACGAGGCGGTGCGAGGCCAGGTGCAACCAACTACTTCTGGATGTG	1261

QY	1261	AGCGGTTGCGTGGAGACCCCCCGCCAGCGTGGAAAGCCCCCGACGATGTCACCCCGCCAGAAAG	1320
Db	1262	AGCGTTGCTGCTGACACCCCGCCAGCTGGAAAGCCCCCGACGATGCCACCGCCAGAAAG	1321
QY	1321	CCCTTATGACCCCTTGCGGCCGCGCTGTGCTTACTGTTACTTGTGAGACCCCTTCTCTCT	1380
Db	1322	CCCTTATGACCCCTTGCGGCCGCGCTGTGCTTACTGTTACTTGTGAGACCCCTTCTCTCT	1381
QY	1381	CCCTGAGGCTGACAGGCTCTGTCCACAGTCTCTGTGGGGGTGGAGAGACTCTTGAGAGA	1440
Db	1382	CCCTGAGGCTGACAGGCTCTGTCCACAGTCTCTGTGGGGGTGGAGAGACTCTTGAGAGA	1441
QY	1441	AGGCTGAGAAAGCTGGAGGCTATGCTTTGGGGGAGACGCCAGAAATGAAGTCTGGTGCAG	1500
Db	1442	AGGCTGAGAAAGCTGGAGGCTATGCTTTGGGGGAGACGCCAGAAATGAAGTCTGGTGCAG	1501
QY	1501	GAGCCCAACTGCTGGCGGCCAGCTGCTATGTAAAGCCCTTACTGTTGTTGAGACCC	1560
Db	1502	GAGCCCAACTGCTGGCGGCCAGCTGCTATGTAAAGCCCTTACTGTTGTTGAGACCC	1561
QY	1561	CACCCCAAGAACCAATATCAATTAAGTAGCATTCGCCAA	1599
Db	1562	CACCCCAAGAACCAATATCAATTAAGTAGCATTCGCCAA	1600

EARLIER APPLICATION NUMBER: 60/048, 892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049, 374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048, 878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070, 923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092, 921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094, 657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 90
LENGTH: 1533
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (123)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1522)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1527)
OTHER INFORMATION: n equals a,t,g, or c
US-10-023-282-90
Query Match 39.0%; Score 624; DB 9; Length 1533;
Best Local Similarity 99.98; Pred. No. 0;
Matches 674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 925 GTGGTCCCGAGAGACTTGTGCTAGTCTGCGACTGTGCACTGCACTGCGACCTGGCCAGT 984

|||||
824 GTGGTCCCGAGAGACTTGTGCTAGTCTGCGACTGTGCACTGCGACCTGGCCAGT 883
QY 985 GAGATGTTGCTGCACCATGGGCGCTGTGCAGTGGCGCTGATGATCTTTACGTG 1044
Db 884 GAGATGTTGCTGCACCATGGGCGCTGTGCAGTGGCGCTGATGATCTTTACGTG 943
QY 1045 CGGGCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104
Db 944 CGGGCGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1003
QY 1105 CATATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1164
Db 1004 CATATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1063
QY 1165 CCCAAGATGGGAAAGTGTGTTGCACTGATGGGAATGATGATGATGATGATGATG 1224
Db 1064 CCCAAGATGGGAAAGTGTGTTGCACTGATGGGAATGATGATGATGATGATGATG 1123
QY 1225 CAGGCGCAGTGCACCCAACTACTTGTGATGTCAGCGGTTGCTGAGAGCCCGCC 1284
Db 1124 CAGGCGCAGTGCACCCAACTACTTGTGATGTCAGCGGTTGCTGAGAGCCCGCC 1183
QY 1285 AGCTGGAAGCCCGAGAGATGCGACCCGAGAGAGCCCTTATGACCCCTGGCGGCT 1344
Db 1184 AGCTGGAAGCCCGAGAGATGCGACCCGAGAGAGCCCTTATGACCCCTGGCGGCT 1243
QY 1345 GTGCTTGTAGTGTACTTGCAGACCCCTTCTCTCCCTTCCCTGAGGCTGAGGCTGTGCA 1404
Db 1244 GTGCTTGTAGTGTACTTGCAGACCCCTTCTCTCCCTTCCCTGAGGCTGAGGCTGTGCA 1303
QY 1405 CAGTCTGTGTGGGTGAGAGACTCTCTGAGAGAGGAGAGAGGAGAGAGGAGAGGAG 1464
Db 1304 CAGTCTGTGTGGGTGAGAGACTCTCTGAGAGAGGAGAGAGGAGAGGAGAGGAG 1363
QY 1465 TTTGGGGGAGAGCCAGAGATGAGTCTGTGGTGCAGAGCCAGCTGGGCGGCACT 1524
Db 1364 TTTGGGGGAGAGCCAGAGATGAGTCTGTGGTGCAGAGCCAGCTGGGCGGCACT 1423
QY 1525 GCTATGTAGGCTTGTGTTGTTGAGACCCCGACCCCGACGAAATCCAAATA 1584
Db 1424 GCTATGTAGGCTTGTGTTGTTGAGACCCCGACCCCGACGAAATCCAAATA 1483
QY 1585 AAGTACATTCCTCAA 1599
Db 1484 AAGTACATTCCTCAA 1498
RESULT 3
US-09-970-516-1
Sequence 1, Application US/09970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: No. US20020099029A1artis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970, 516
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1155
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1155)
OTHER INFORMATION:
US-09-970-516-1
Query Match 27.1%; Score 433; DB 10; Length 1155;
Best Local Similarity 100.0%; Pred. No. 1.2e-215;


```

; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1756
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1756

Query Match          11.9%; Score 191; DB 10; Length 394;
Best Local Similarity 99.3%; Pred. No. 2.4e-89;
Matches 291; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1210 GTTAGCGAGCGCGTGCAGGCGCCAGGTGCAACCAACTCTTGAGTGTGAGCGGTTGC 1269
    |||||||
DB 394 GTTAGCGAGCGCGTGCAGGCGCCAGGTGCAACCAACTCTTGAGTGTGAGCGGTTGC 335
    |||||||
QY 1270 GTGAGAGCGCGCGCGTGAAGCCAGAGAGTCCAGCGCCAGAGAGCGGTTATGA 1329
    |||||||
DB 334 GTGAGAGCGCGCGCGTGAAGCCAGAGAGTCCAGCGCCAGAGAGCGGTTATGA 275
    |||||||
QY 1330 CCCCTGGGCGCGCGTGTGCTTACTTGTCTACTTGTGAGAGCGCTTCTCCCTAGAGGC 1389
    |||||||
DB 274 CCCCTGGGCGCGCGTGTGCTTACTTGTCTACTTGTGAGAGCGCTTCTCCCTAGAGGC 215
    |||||||
QY 1390 TGCAGGCGCTTCCACAGCTCTGTGGGGGTGGAGAGACTTCTTGGAGAGGTTGGA 1449
    |||||||
DB 214 TGCAGGCGCTTCCACAGCTCTGTGGGGGTGGAGAGACTTCTTGGAGAGGTTGGA 155
    |||||||
QY 1450 AGGTGAGAGCTATGCTTGGGGGGGAGAGCCAGAGATGAAGTCCTGAGGTA 1502
    |||||||
DB 154 AGGTGAGAGCTATGCTTGGGGGGGAGAGCCAGAGATGAAGTCCTGAGGTA 102
    |||||||

RESULT 7
US-09-796-692-2905
; Sequence 2905, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077,001200
; CURRENT FILING DATE: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
```

```

; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 2905
; LENGTH: 199
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-2905

Query Match          10.9%; Score 175; DB 9; Length 199;
Best Local Similarity 100.0%; Pred. No. 5.8e-81;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1088 CCATGAGAGAGGCGGCGATATGAGTGAATGCCCTACTGTGATATGATGCCCGTGG 1147
    |||||||
DB 1 CCATGAGAGAGGCGGCGATATGAGTGAATGCCCTACTGTGATATGATGCCCGTGG 60
    |||||||
QY 1148 TCGCCTCCGCTTGGAGCCCAAGAGTGGAAAGTGTGTTGCACTGATGGAGATTGA 1207
    |||||||
DB 61 TCGCCTCCGCTTGGAGCCCAAGAGTGGAAAGTGTGTTGCACTGATGGAGATTGA 120
    |||||||
QY 1208 TGTGTTAGCGAGCGCGTGCAGGCGCAGGTGCACCCAACTACTTCTGTGATGTCAG 1262
    |||||||
DB 121 TGTGTTAGCGAGCGCGTGCAGGCGCAGGTGCACCCAACTACTTCTGTGATGTCAG 175
    |||||||

RESULT 8
US-10-040-862-2905
; Sequence 2905, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE REFERENCE: 014058-013520US
; CURRENT FILING DATE: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
```

```

NAME/KEY: misc feature
LOCATION: (322)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (337)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (350)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (356)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (367)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (382)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (384)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (386)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (391)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (400)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (412)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (451)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9248

Query Match      3.1%; Score 49; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 2,9e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      311  AATCTCCCTTCACGCTGATGCTCAGTCAGCGGCGGAGCACGCGCGGGA 359
DB      10  AATCTCCTTCACGCTGATGCTCAGTCAGCGGCGGAGCACGCGCGGGA 58
|||||
|||||

RESULT 10
US-09-796-487-17/c
: Sequence 17, Application US/09796487
: Patent No. US20020042358A1
: GENERAL INFORMATION:
: APPLICANT: Spiegel, Sarah
: TITLE OF INVENTION: Splicing Kinase, Cloning, Expression and Methods of Use
: FILE REFERENCE: 07320001aa (203957-0001)
: CURRENT APPLICATION NUMBER: US/09/796,487
: CURRENT FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: US 60/186,532
: PRIOR FILING DATE: 2000-03-03
: PRIOR APPLICATION NUMBER: US 09/530,868
: PRIOR FILING DATE: 2000-05-05
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 17
: LENGTH: 30
: TYPE: DNA
: ORGANISM: Unknown
: FEATURE:
: OTHER INFORMATION: Artificial sequence of the antisense primer sphk1-GSP5 in line
: NAME/KEY: misc_feature
: LOCATION: (1)..(30)
: OTHER INFORMATION: Corresponding to the artificial sequence of antisense primer sphk1-GSP5 in line

```


OTHER INFORMATION: -GSP5 in line 5-6, page 8.
US-09-796-487-17

Query Match 1.9%; Score 30; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1300 CAGATGCCACGCCAGAGAGCCTTATGA 1329
|||||
DB 30 CAGATGCCACGCCAGAGAGCCTTATGA 1

RESULT 11
US-09-970-516-5
Sequence 5, Application US/09970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: NO. US20020099029A1artis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970,516
CURRENT FILING DATE: 2001-10-04
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1149
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1149)
OTHER INFORMATION:
US-09-970-516-5

Query Match 1.6%; Score 25; DB 10; Length 1149;
Best Local Similarity 100.0%; Pred. No. 0.0088;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1068 GCTGCTGGCCTCTTCTGCGCATG 1092
|||||
DB 891 GCTGCTGGCCTCTTCTGCGCATG 915

RESULT 12
US-09-784-810A-3
Sequence 3, Application US/09784810A
Patent No. US20020082203A1
GENERAL INFORMATION:
APPLICANT: RASTELT, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
FILE REFERENCE: 10716-08
CURRENT APPLICATION NUMBER: US/09/784,810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 1759
TYPE: DNA
ORGANISM: Mus musculus
US-09-784-810A-3

Query Match 1.6%; Score 25; DB 10; Length 1759;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1068 GCTGCTGGCCTCTTCTGCGCATG 1092
|||||

DB 1217 GCTGCTGGCCTCTTCTGCGCATG 1241

RESULT 13
US-09-918-995-26954
Sequence 26954, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 26954
LENGTH: 459
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(459)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-26954

Query Match 1.4%; Score 22; DB 9; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 977 TGCCACGAGAGATGTTGCTGC 998
|||||
DB 295 TGCCACGAGAGATGTTGCTGC 316

RESULT 14
US-09-796-487-16
Sequence 16, Application US/09796487
Patent No. US20020042358A1
GENERAL INFORMATION:
APPLICANT: Spiegel, Sarah
TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
FILE REFERENCE: 07320001a (7033957-0001)
CURRENT APPLICATION NUMBER: US/09/796,487
CURRENT FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: US 60/186,532
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: US 09/530,868
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 27
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Artificial sequence of sense primer sphk1-GSP4, shown in line 4
OTHER INFORMATION: page 8
NAME/KEY: misc-feature
LOCATION: (1)..(27)
OTHER INFORMATION: Corresponding to artificial sequence of the sense primer sphk1-
US-09-796-487-16

Query Match 1.3%; Score 21; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 175 ATGATCCAGCGCGCGGCCCC 195
|||||
DB 7 ATGATCCAGCGCGCGGCCCC 27

Mon Jul 14 09:16:27 2003

us-09-784-810a-1_copy_2_1600.olg.rnpb

```
RESULT 15
US-09-796-487-15/c
; Sequence 15, Application US/09796487
; Patent No. US20020042358A1
; GENERAL INFORMATION:
; APPLICANT: Spiegel, Sarah
; TITLE OF INVENTION: Sphingosin Kinase, Cloning, Expression and Methods of Use
; FILE REFERENCE: 07320001aa (2033957-0001)
; CURRENT APPLICATION NUMBER: US/09/796,487
; CURRENT FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/186,532
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/530,868
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Artificial sequence of antisense primer hspk1-GSP, shown in line
; OTHER INFORMATION: 34, page 7.
; NAME/KEY: misc.feature
; LOCATION: (1)..(20)
; OTHER INFORMATION: Corresponding to the sequence of antisense primer hspk1-GSP3, shc
; OTHER INFORMATION: wn in line 34, page 7.
US-09-796-487-15

Query Match          1.3%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      761 CCTTCTGCGCTCTGCGAGCC 780
        |||
Db      20 CCTTCTGCGCTCTGCGAGCC 1

Search completed: July 12, 2003, 18:06:39
Job time : 278 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 15:59:52 ; Search time 83 Seconds

(without alignments)
5908.146 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599
Sequence: 1 cccccggggcctctatagc.....aataaagtgcattcccaa 1599

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PC10S.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Dacfilles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21	1.3	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
C 2	21	1.3	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
C 3	20	1.3	10952	1 US-08-602-036A-1	Sequence 1, Appl1
C 4	20	1.3	10952	2 US-08-502-374A-1	Sequence 1, Appl1
C 5	20	1.3	10952	2 US-08-642-407A-1	Sequence 1, Appl1
C 6	19	1.2	396	1 US-07-924-063A-4	Sequence 4, Appl1
C 7	19	1.2	4285	4 US-09-040-774-1	Sequence 1, Appl1
C 8	19	1.2	10079	2 US-08-476-866-20	Sequence 20, Appl1
C 9	18	1.1	1260	3 US-08-789-982-1	Sequence 1, Appl1
C 10	18	1.1	1722	3 US-08-691-563C-58	Sequence 58, Appl1
C 11	18	1.1	1761	3 US-08-861-747-1	Sequence 1, Appl1
C 12	18	1.1	1889	3 US-08-861-747-3	Sequence 1, Appl1
C 13	18	1.1	2081	3 US-08-935-855-21	Sequence 21, Appl1
C 14	18	1.1	2834	4 US-09-305-384-6	Sequence 6, Appl1
C 15	18	1.1	6235	4 US-09-305-384-5	Sequence 5, Appl1
C 16	18	1.1	6679	4 US-09-305-384-1	Sequence 1, Appl1
C 17	18	1.1	9704	4 US-09-814-951A-3	Sequence 3, Appl1
C 18	18	1.1	29	1 US-08-293-086-8	Sequence 8, Appl1
C 19	17	1.1	29	1 US-08-244-993-8	Sequence 8, Appl1
C 20	17	1.1	29	2 US-08-861-306-8	Sequence 8, Appl1
C 21	17	1.1	29	2 US-08-468-037A-35	Sequence 35, Appl1
C 22	17	1.1	29	2 US-08-471-973A-35	Sequence 35, Appl1
C 23	17	1.1	29	2 US-08-465-880-16	Sequence 16, Appl1
C 24	17	1.1	29	3 US-09-035-357-35	Sequence 35, Appl1
C 25	17	1.1	29	3 US-09-016-520-17	Sequence 17, Appl1
C 26	17	1.1	29	3 US-09-144-611-8	Sequence 8, Appl1
C 27	17	1.1	29	4 US-09-130-973-17	Sequence 17, Appl1

28	17	1.1	29	4	US-09-477-902-17	Sequence 17, Appl1
29	17	1.1	29	4	US-09-453-514A-8	Sequence 8, Appl1
30	17	1.1	29	4	US-09-135-202-35	Sequence 35, Appl1
31	17	1.1	29	4	US-08-802-331-16	Sequence 16, Appl1
32	17	1.1	147	4	US-09-536-094-2	Sequence 2, Appl1
33	17	1.1	255	3	US-08-806-326-6	Sequence 6, Appl1
34	17	1.1	356	3	US-08-806-326-7	Sequence 7, Appl1
35	17	1.1	356	4	US-09-536-094-3	Sequence 3, Appl1
36	17	1.1	433	1	US-08-428-733A-1	Sequence 1, Appl1
C 37	17	1.1	843	1	US-09-427-700-11	Sequence 11, Appl1
C 38	17	1.1	969	2	US-08-700-637-1	Sequence 1, Appl1
C 39	17	1.1	1008	4	US-09-134-001C-2791	Sequence 2791, Ap
C 40	17	1.1	1149	2	US-08-466-103A-3	Sequence 3, Appl1
C 41	17	1.1	1241	4	US-09-149-476-220	Sequence 220, App
C 42	17	1.1	1289	2	US-08-344-833-1	Sequence 1, Appl1
C 43	17	1.1	1398	2	US-08-896-320-2	Sequence 2, Appl1
C 44	17	1.1	1442	1	US-08-152-483B-8	Sequence 8, Appl1
C 45	17	1.1	1461	3	US-08-889-841B-7	Sequence 7, Appl1

ALIGNMENTS

```
RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CCG 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match
Best Local Similarity 1.3%; Score 21; DB 4; Length 4403765;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 TGTGCTCAGCCTGCGCTGGGG 689
Db 1033190 TGTGCTCAGCCTGCGCTGGGG 1033170

RESULT 2
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
```

SOFTWARE: Patent Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37RV
US-09-103-840A-1

Query Match 1.3%; Score 21; DB 4; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 669 TGCTCTAGCCTGCGCTGGG 689
DB 1033188 TGTGCTAGCCTGCGCTGGG 1033168

RESULT 3
US-08-602-036A-1
Sequence 1, Application US/08602036A
Patent No. 5789248
GENERAL INFORMATION:
APPLICANT: Oeystein, Fodstad
APPLICANT: Hovig, Elvind
APPLICANT: Engebraaten, Olav
APPLICANT: Maelandmo, Gunhild H.
TITLE OF INVENTION: CAPL SPECIFIC OLIGONUCLEOTIDES AND
METHODS OF INHIBITING METASTATIC CANCER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,036A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kerner, Ann-Louise
REGISTRATION NUMBER: 33,523
REFERENCE/DOCKET NUMBER: HYZ-039CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10952 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-602-036A-1

Query Match 1.3%; Score 20; DB 1; Length 10952;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 GGAGACTCCTCTGGAGAAG 1443
DB 10557 GGAGACTCCTCTGGAGAAG 10576

RESULT 4
US-08-502-374A-1
Sequence 1, Application US/08502374A
Patent No. 5872007
GENERAL INFORMATION:
APPLICANT: Fodstad, Oeystein
APPLICANT: Hovig, Elvind
APPLICANT: Engebraaten, Olav
APPLICANT: Maelandmo, Gunhild H.
TITLE OF INVENTION: CAPL-SPECIFIC OLIGONUCLEOTIDES AND
METHODS OF INHIBITING METASTATIC CANCER
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/502,374A
FILING DATE: 14-Jul-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kerner, Ann-Louise
REGISTRATION NUMBER: 33,523
REFERENCE/DOCKET NUMBER: HYZ-039DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-526-6000
TELEFAX: 617-526-5000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10952 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-502-374A-1

Query Match 1.3%; Score 20; DB 2; Length 10952;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 GGAGACTCCTCTGGAGAAG 1443
DB 10557 GGAGACTCCTCTGGAGAAG 10576

RESULT 5
US-08-642-407A-1
Sequence 1, Application US/08642407A
Patent No. 5877308
GENERAL INFORMATION:
APPLICANT: Oeystein, Fodstad
APPLICANT: Hovig, Elvind
APPLICANT: Engebraaten, Olav
APPLICANT: Maelandmo, Gunhild H.
TITLE OF INVENTION: CAPL SPECIFIC OLIGONUCLEOTIDES AND
METHODS OF INHIBITING METASTATIC CANCER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston

STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,407A
FILING DATE: 03-May-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Keirner, Ann-Louise
REGISTRATION NUMBER: 33,523
REFERENCE/DOCKET NUMBER: HTZ-039CPDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10952 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-642-407A-1

Query Match 1.3%; Score 20; DB 2; Length 10952;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1424 GGAGACTCTCTGGAGAAG 1443
|||||
DB 10557 GGAGACTCTCTGGAGAAG 10576

RESULT 6
US-07-924-063A-4/C
Sequence 4, Application US/07924063A
Patent No. 5489514
GENERAL INFORMATION:
APPLICANT: TSUJI, SHOJI; MIYATAKE, TADASHI; UCHIDA, YOKO;
APPLICANT: IHARA, YASUO
TITLE OF INVENTION: DNA CODING FOR GROWTH-INHIBITORY FACTOR AND
TITLE OF INVENTION: USE THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIERMAN & MUSERLIAN
STREET: 600 THIRD AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/924,063A
FILING DATE: 19920828
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/01714
FILING DATE: 13-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 410165/1990
FILING DATE: 13-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: CHARLES A. MUSERLIAN

REGISTRATION NUMBER: 19,683
REFERENCE/DOCKET NUMBER: TSU-23
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-661-8000
TELEFAX: 212-661-8002
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 bp
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE: brain
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: GIF gene and flanking sequence
US-07-924-063A-4

Query Match 1.2%; Score 19; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 906 GGTGCTTCTCAGTCGACG 924
|||||
DB 268 GGTGCTTCTCAGTCGACG 250

RESULT 7
US-09-040-774-1
Sequence 1, Application US/09040774
Patent No. 6207811
GENERAL INFORMATION:
APPLICANT: Trygsvason, Karl
APPLICANT: Kestila, Marjo
APPLICANT: Leukeril, Ulla
APPLICANT: Mannikko, Minna
TITLE OF INVENTION: Nephlin Gene and Protein
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, Suite 3200
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,774
FILING DATE: 18 MAR 1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chao, Mark
REGISTRATION NUMBER: 37,293
REFERENCE/DOCKET NUMBER: 97,842
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)913-0001
TELEFAX: (312)913-0002

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 1..56
OTHER INFORMATION: /note= "putative signal peptide"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3723
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 67..3723
FEATURE:
NAME/KEY: misc_feature
LOCATION: 121..122
OTHER INFORMATION: /note= "deletion mutation"
OTHER INFORMATION: /note= "deletion mutation"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3800..3804
OTHER INFORMATION: /note= "nonsense mutation in exon"
OTHER INFORMATION: 26 FIN-Minor"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3178..3258
OTHER INFORMATION: /note= "putative transmembrane"
OTHER INFORMATION: domain"
US-09-040-774-1

Query Match 1.2%; Score 19; DB 4; Length 4285;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 TCCAGCGGGCGGCCCGG 198
DB 2067 TCCAGCGGGCGGCCCGG 2085

RESULT 8
US-08-476-866-20/c
Sequence 20, Application US/08476866
Patent No. 5994339
GENERAL INFORMATION:
APPLICANT: CRAPO, JAMES D.
APPLICANT: PRIDOVICH, IRVIN
APPLICANT: CURRY, TIM
APPLICANT: DAY, BRIAN J.
APPLICANT: FOLZ, RODNEY J.
APPLICANT: FREEMAN, BROCE A.
TITLE OF INVENTION: SUPEROXIDE DISMUTASE AND MIMETICS THEREOF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476, 866
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/322,766
FILING DATE: 13-OCT-1994
APPLICATION NUMBER: US 08/136,207
FILING DATE: 15-OCT-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-74
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 10079 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 5086..5803
US-08-476-866-20

Query Match 1.2%; Score 19; DB 2; Length 10079;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 TAAGAGCTGAGGAGGA 105
DB 2891 TAAGAGCTGAGGAGGA 2873

RESULT 9
US-08-789-982-1
Sequence 1, Application US/08789982
Patent No. 6037146
GENERAL INFORMATION:
APPLICANT: Sathya, Ganesh
APPLICANT: Bergsma, Derek
TITLE OF INVENTION: CDNA CLONE HERC90 THAT ENCODES
TITLE OF INVENTION: A NOVEL 7- TRANSMEMBRANE RECEPTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789, 982
FILING DATE: 28-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T.
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50050
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4026
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1260 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-789-982-1

Query Match 1.1%; Score 18; DB 3; Length 1260;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 GCGTCTGCTGCTGCTGA 238
|||||
DB 260 GCGTCTGCTGCTGCTGA 277

RESULT 10
US-08-691-563C-58/C
Sequence 58, Application US/08691563C
Patent No. 6001987

GENERAL INFORMATION:
APPLICANT: HERVE PERRON
APPLICANT: FREDERIC BESEME
APPLICANT: FREDERIC BEDIN
APPLICANT: GLAUCIA PARANHOS-BACCALA
APPLICANT: FLORENCE KOMURIAN-PRADEL
APPLICANT: COLETTE JOLIVET
APPLICANT: BERNARD MANDRAND
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
TITLE OF INVENTION: THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/691,563C
FILING DATE: 02-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 38588
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:

LENGTH: 1722 base pairs

TYPE: nucleotide

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-691-563C-58

Query Match 1.1%; Score 18; DB 3; Length 1722;
Best Local Similarity 100.0%; Pred. No. 41;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1497 TCAGAGCCAGCTGCT 1514
|||||
DB 1658 TCAGAGCCAGCTGCT 1641

RESULT 11

US-08-861-747-1
Sequence 1, Application US/08861747
Patent No. 6020158

GENERAL INFORMATION:

APPLICANT: MUNROE, Donald G.

APPLICANT: VYAS, Tejal B.

TITLE OF INVENTION: A HUMAN EDG-6 RECEPTOR HOMOLOG

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIKALDO, Marmelstein, Murray & Oram LLP

STREET: 655 15th St., NW, Suite 330 - G Street Lobby

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-5701

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,747
FILING DATE: 22-MAY-1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Jahns, Kristina M.

REGISTRATION NUMBER: 41,092

REFERENCE/DOCKET NUMBER: P8074-7003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 638-5000

TELEFAX: (202) 638-4810

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1761 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-861-747-1

Query Match 1.1%; Score 18; DB 3; Length 1761;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 GCGTCTGCTGCTGCTGA 238
|||||
DB 218 GCGTCTGCTGCTGCTGA 235

RESULT 12

US-08-861-747-3

Sequence 3, Application US/08861747

Patent No. 6020158

GENERAL INFORMATION:

APPLICANT: MUNROE, Donald G.

APPLICANT: VYAS, Tejal B.

TITLE OF INVENTION: A HUMAN EDG-6 RECEPTOR HOMOLOG

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIKALDO, Marmelstein, Murray & Oram LLP

STREET: 655 15th St., NW, Suite 330 - G Street Lobby

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-5701

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,747
FILING DATE: 22-MAY-1997

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Jahns, Kristina M.
REGISTRATION NUMBER: 41,092
REFERENCE/DOCKET NUMBER: P8074-7003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-5000
TELEFAX: (202) 638-4810
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1889 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-861-747-3

Query Match 1.1%; Score 18; DB 3; Length 1889;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 GCGTGTGCTGCTGCTGA 238
DB 304 GCGTGTGCTGCTGCTGA 321

RESULT 13
US-08-935-855-21/C
Sequence 21, Application US/08935855
Patent No. 6066485
GENERAL INFORMATION:
APPLICANT: Guttridge, Mark
TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,855
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-002 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 2081 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-935-855-21

Query Match 1.1%; Score 18; DB 3; Length 2081;

Best Local Similarity 100.0%; Pred. No. 41;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 GCGAGCGGAGCCGCGGG 166
DB 97 GCGAGCGGAGCCGCGGG 80

RESULT 14
US-09-305-384-6/C
Sequence 6, Application US/09305384
Patent No. 6242218
GENERAL INFORMATION:
APPLICANT: Treco, Douglas A.
APPLICANT: Heitlein, Michael W.
TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
FILE REFERENCE: 07236/017001
CURRENT APPLICATION NUMBER: US/09/305,384
CURRENT FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,649
EARLIER FILING DATE: 1998-05-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 2834
TYPE: DNA
ORGANISM: Homo sapiens
US-09-305-384-6

Query Match 1.1%; Score 18; DB 4; Length 2834;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1497 TCAGAGCCACAGCTGGCT 1514
DB 633 TCAGAGCCACAGCTGGCT 616

RESULT 15
US-09-305-384-5/C
Sequence 5, Application US/09305384
Patent No. 6242218
GENERAL INFORMATION:
APPLICANT: Treco, Douglas A.
APPLICANT: Heitlein, Michael W.
TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
FILE REFERENCE: 07236/017001
CURRENT APPLICATION NUMBER: US/09/305,384
CURRENT FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,649
EARLIER FILING DATE: 1998-05-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 6235
TYPE: DNA
ORGANISM: Homo sapiens
US-09-305-384-5

Query Match 1.1%; Score 18; DB 4; Length 6235;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1497 TCAGAGCCACAGCTGGCT 1514
DB 2522 TCAGAGCCACAGCTGGCT 2505

Search completed: July 12, 2003, 18:01:58
Job time : 91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 15:19:57 ; Search time 2386 Seconds

(without alignments)
10853.555 Million cell updates/sec

Title: US-09-784-810a-1_COPY_2_1600

Perfect score: 1599

Sequence: 1 cccccggggtctctatagc.....aataaagtgcattcccaa 1599

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapept 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	608	38.0	662	13	BM386957 UI-R-CNI-
2	590	36.9	2133	11	BC014439 Homo sapi
3	537	33.6	1005	13	BM559257 AGENCOURT
4	527	33.0	676	10	AM963415
5	467	29.2	720	14	BQ109456
6	447	28.0	772	9	AI769914 wj30d06.x

Result No.	Score	Query Match	Length	DB ID	Description
7	440	27.5	447	9	A1042283
8	440	27.5	587	9	AU147968
9	422	26.4	852	10	BE274434
10	422	26.4	1075	13	BM557357
11	405	25.3	618	10	AM662445
12	399	25.0	416	13	BM471470
13	393	24.6	603	10	BE275818
14	391	24.5	1115	13	BM550039
15	391	24.5	1134	14	BM808666
16	381	23.8	675	9	AU134361
17	380	23.8	388	9	A1074186
18	378	23.6	702	10	BE464487
19	374	23.4	638	9	A1972156
20	360	22.5	478	12	BE740866
21	337	21.1	1232	13	BM556915
22	336	21.0	355	14	BM027486
23	322	20.1	697	13	BI255900
24	321	20.1	409	10	AM192475
25	312	19.5	999	14	BM675531
26	297	18.6	348	14	BM709951
27	294	18.4	809	12	BM280830
28	293	18.3	793	12	BM678689
29	291	18.2	1100	14	BM916526
30	291	18.2	1137	14	BM810136
31	290	18.1	888	14	BM647377
32	288	18.0	604	9	AU122376
33	282	17.6	856	13	BI757316
34	277	17.3	534	12	BE904632
35	276	17.3	1054	14	BM808698
36	265	16.6	412	9	AA639414
37	258	16.1	383	12	BM057227
38	241	15.1	650	12	BM104264
39	240	15.0	271	14	D31133
40	220	14.4	316	9	AA372241
41	220	13.8	973	10	BE273426
42	206	12.9	266	13	BM471402
43	198	12.4	318	13	BI013436
44	198	12.4	732	12	BE728564
45	194	12.1	413	13	BI013566

ALIGNMENTS

RESULT 1
LOCUS BM386957/c
DEFINITION UI-R-CNI-cj1-c-13-0-UI.s1 UI-R-CNI Rattus norvegicus CDNA clone
ACCESSION BM386957
VERSION BM386957.1 GI:18187010
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 662)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477

COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: mssoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A

Db 84 TAAGCCCTTAGTTTGTGTAGAGACCCCAACCAACCAATTAAGTGAC 25
 QY 1592 ATTCACAA 1599
 Db 24 ATTCACAA 17

RESULT 2
 LOCUS BC014439 2133 bp mRNA linear HTC 19-DEC-2001
 DEFINITION Homo sapiens, clone IMAGE:4871343, mRNA.
 ACCESSION BC014439
 VERSION BC014439.1 GI:17939597
 KEYWORDS HTC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Strausberg, R.
 Direct Submission
 Submitted (17-SEP-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letlicia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Nees, Pawan Pandoh, Anna-Ilisa Prabhu, Parvaneh Seedei, Jacqueline
 Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Zaira.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 34 Row: 1 Column: 9
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 11464966
 This clone has the following problem: incomplete processing.
 Location/Qualifiers
 1. 2133
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4871343"
 /tissue_type="Skin, melanotic melanoma, high MDR."
 /clone_id="NIH_MGC_49"
 /lab_host="DH10B-R"
 /note="Vector: pORB7"
 BASE COUNT 350 a 681 c 719 g 383 t
 ORIGIN

Query Match 36.9%; Score 590; DB 11; Length 2133;
 Best Local Similarity 100.0%; Pred. No. 8.4e-272;
 Matches 590; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 AGCCAGGCTCCGGGGGGAAGGCGAGCCCAAGCGGCTCGAGCGCCGCTGGG 77
 Db 536 AGCCAGGCTCCGGGGGGAAGGCGAGCCCAAGCGGCTCGAGCGCCGCTGGG 595
 QY 78 CAGCAGCATAGAGAGCTGAAGGCGAGCGCGCCCAAGGCGCGCCCAAGCGCC 137
 Db 596 CAGCAGCATAGAGAGCTGAAGGCGAGCGCGCCCAAGGCGCGCCCAAGCGCC 655

QY 138 AGGACCCCTTGCGAGCGGAGCGCGGGCTGAGATTATGATCCAGCGGGCGCCCG 197
 Db 656 AGGACCCCTTGCGAGCGGAGCGCGGGCTGAGATTATGATCCAGCGGGCGCCCG 715
 QY 198 GGGCGTGTCCCGGGGCGGCGCTGCGGCTGCTGCTGCTGAACCCGCGGCGAAGG 257
 Db 716 GGGCGTGTCCCGGGGCGGCGCTGCGGCTGCTGCTGCTGAACCCGCGGCGAAGG 775
 QY 258 CAAGGCTTGAGAGCTCTCCGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 317
 Db 776 CAAGGCTTGAGAGCTCTCCGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 835
 QY 318 CTTCAGCTGATGCTCACTAGAGCGCGGAGACACGCGGAGCTGTCGCGTGAAGA 377
 Db 836 CTTCAGCTGATGCTCACTAGAGCGCGGAGACACGCGGAGCTGTCGCGTGAAGA 895
 QY 378 GCTGGGCGGCTGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
 Db 896 GCTGGGCGGCTGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
 QY 438 GAACGGCTCATGAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 497
 Db 956 GAACGGCTCATGAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
 QY 498 CCACAGAGCTCTGCGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
 Db 1016 CCACAGAGCTCTGCGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075
 QY 558 GGTCCCAATGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607
 Db 1076 GGTCCCAATGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1125

RESULT 3
 BM559257 1005 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT_6598353 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5474413
 DEFINITION 5', mRNA sequence.
 ACCESSION BM559257
 VERSION BM559257.1 GI:18802694
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC <http://mgc.ncl.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LHC1987 row: 3 column: 14
 High quality sequence stop: 618.
 Location/Qualifiers
 1. 1005
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5474413"
 /clone_id="NIH_MGC_41"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORB7; Site 1: XhoI; Site 2:
 EcoRI/XhoI sites using the following 5' adaptor:
 GCCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
SuperScript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 157 a 333 c 348 g 167 t
Query Match 33.6%; Score 537; DB 13; Length 1005;
Best Local Similarity 100.0%; Pred. No. 2e-246;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

18 AGCCAGGCTCCGGGCGGAGAGCCGAGCCGCTGAGAGCCGCGCTGG 77
169 AGCCAGGCTCCGGGCGGAGAGCCGAGCCGCTGAGAGCCGCGCTGG 228
78 CAGCAGCGATAGAGAGCTGAGAGAGCCGCGAGAGAGAGAGAGAGAG 137
229 CAGCAGCGATAGAGAGCTGAGAGAGCCGCGAGAGAGAGAGAGAGAG 288
138 AGGAGCCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 197
289 AGGAGCCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
198 GGGGCTGCTCCCGGCGGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 257
349 GGGGCTGCTCCCGGCGGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 408
258 CAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 317
409 CAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 468
318 CTTACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377
469 CTTACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
378 GCTGGGCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 437
529 GCTGGGCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588
438 GAAGGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 497
589 GAAGGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
498 CCCAGCAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 554
649 CCCAGCAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 705

RESULT 4
AM963415 676 bp mRNA linear EST 01-JUN-2000
LOCUS EST375488 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.
DEFINITION AM963415
VERSION AM963415.1 GI:8153251
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 676)
Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C., Holt
I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and
Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johngett@igf.org
Plate: 187
Seq primer: Reverse.

FEATURES
source Location/Qualifiers
1..676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGH"
/note="Vector: pBluescriptPSKm"

BASE COUNT 147 a 220 c 188 g 121 t

Query Match 33.0%; Score 527; DB 10; Length 676;
Best Local Similarity 100.0%; Pred. No. 1.2e-241;
Matches 527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1070 TGTGCGGCTCTTCTGCGCCATGAGAGAGAGAGAGAGAGAGAGAGAG 1129
527 TGTGCGGCTCTTCTGCGCCATGAGAGAGAGAGAGAGAGAGAGAGAG 468
1130 TGTATATATGCGCCGCTGCGCTTCCGCTGAGAGAGAGAGAGAGAG 1189
467 TGTATATATGCGCCGCTGCGCTTCCGCTGAGAGAGAGAGAGAGAG 408
1190 CAGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1249
407 CAGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
1250 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1309
347 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 288
1310 CGCCAG 1369
287 CGCCAG 228
1370 CTTCT 1429
227 CTTCT 168
1430 TCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1489
167 TCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108
1490 TCTGAG 1549
107 TCTGAG 48
1550 TCTGAG 1596
47 TCTGAG 1

RESULT 5
BO109456 720 bp mRNA linear EST 16-APR-2002
LOCUS IMAGE5217439 5' mRNA sequence.
DEFINITION IMAGE5217439 5' mRNA sequence.
VERSION BO109456
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 720)
Kale, P. I., Harsch, T. J., Folta, P. A., Nelson, D. O., Sanders, C. G. and
Prange, C. K.
TITLE The I.M.A.G.E. Consortium quality control effort: clone
ressequencing for verification
JOURNAL Unpublished (2001)
COMMENT Other ESTs: B1912225
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA

Email: help@image.llnl.gov

This read has been verified (found to hit its original self in the correct orientation), as part of the I.M.A.G.E. Consortium quality control effort. High quality sequence is having 100 or more base pairs with a phred quality value of 20 or greater, whereas a sliding window of 4 base pairs with a phred quality value of 15 or greater marks the beginning and end of the sequence. For information on obtaining this clone, please contact info@image.llnl.gov.

BASE COUNT ORIGIN	114 a	239 c	265 g	102 t
----------------------	-------	-------	-------	-------

FEATURES	JOURNAL COMMENT	REFERENCE AUTHORS TITLE	ORGANISM SOURCE	ACCESSION VERSION	LOCUS DEFINITION	RESULT 6
772 bp	EST 21-DEC-1999	NCI-CGAP Kd12 Homo sapiens CDNA clone IMAGE:2404331 3'	human.	AI769914	AI769914	AI769914/c
mRNA	linear	similar to TR:088886 088886 SPHINGOSINE KINASE. ; , mRNA sequence.	human.	AI769914	AI769914	AI769914/c
GI:5236423		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	human.	AI769914	AI769914	AI769914/c
		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	human.	AI769914	AI769914	AI769914/c
		Tumor Gene Index	human.	AI769914	AI769914	AI769914/c
		Unpublished (1997)	human.	AI769914	AI769914	AI769914/c
		Contact: Robert Strausberg, Ph.D.	human.	AI769914	AI769914	AI769914/c
		Email: cgapsb@email.nih.gov	human.	AI769914	AI769914	AI769914/c
		Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.	human.	AI769914	AI769914	AI769914/c
		Emmert-Buck, M.D., Ph.D.	human.	AI769914	AI769914	AI769914/c
		CDNA Library Preparation: M. Bento Soares, Ph.D.	human.	AI769914	AI769914	AI769914/c
		CDNA Library Arrayed by: Greg Lennon, Ph.D.	human.	AI769914	AI769914	AI769914/c
		DNA Sequencing by: Washington University Genome Sequencing Center	human.	AI769914	AI769914	AI769914/c
		Clone distribution: NCI-CGAP clone distribution information can be	human.	AI769914	AI769914	AI769914/c
		found through the I.M.A.G.E. Consortium/ILNI, at:	human.	AI769914	AI769914	AI769914/c
		www-bio.llnl.gov/bdrrp/image/image.html	human.	AI769914	AI769914	AI769914/c
		Insert length: 1211 Std Error: 0.00	human.	AI769914	AI769914	AI769914/c
		Seq primer: -400P from Gibco	human.	AI769914	AI769914	AI769914/c
		High quality sequence stop: 467.	human.	AI769914	AI769914	AI769914/c
		Location/Qualifiers	human.	AI769914	AI769914	AI769914/c

High quality sequence stop: 467.
Location/Qualifiers

High quality sequence stop: 467.
Location/Qualifiers

Db 420 GTCTGTTGACATGAGGGAATGATGATTAGCAGAGCCGTCAGAGGCCAGGTGCACC 361
 QY 1241 CAAACTACTTGTGATGATGACGGTGTGCTGAGACCCCGCCAGCTGAGAAAGCCCAAC 1300
 Db 360 CAAACTACTTGTGATGATGACGGTGTGCTGAGACCCCGCCAGCTGAGAAAGCCCAAC 301
 QY 1301 AGATGCCACGCCAGAAAGAGCCCTTATGACCCCTGAGGCGCGGCTGTGCTTATGATCTAC 1360
 Db 300 AGATGCCACGCCAGAAAGAGCCCTTATGACCCCTGAGGCGCGGCTGTGCTTATGATCTAC 241
 QY 1361 TTGCAGAGACCTTCT 1420
 Db 240 TTGCAGAGACCTTCT 181
 QY 1421 GGAGAGAGCT 1480
 Db 180 GGAGAGAGCT 121
 QY 1481 AGATGAAGTCT 1540
 Db 120 AGATGAAGTCT 61

RESULT 7
 A1042283/c 447 bp mRNA linear EST 24-SEP-1998
 LOCUS oyl1e09.x1 Soares_senseant_fibroblasts_Nbhsf Homo sapiens cDNA
 DEFINITION clone IMAGE:165736 3', mRNA sequence.
 VERSION A1042283
 ACCESSION A1042283.1 GI:3281477
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 447)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 COMMENT Contact: Robert Strausberg, Ph.D.
 Unpublished (1997)

FEATURES
 SOURCE
 1..447
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:165736"
 /clone_lib="Soares_senseant_fibroblasts_Nbhsf"
 /tissue_type="senseant_fibroblast"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Vector: pT73D (Pharmacia) with a modified
 polylinker V-type: phagemid; Site_1: Not I; Site_2: Eco RI
 ; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer [5'
 TGTACCAATCTGAGTGGAGCGCCGATTTTTTTTTTTTTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldi."

BASE COUNT 87 a 142 c 122 g 97 t
 ORIGIN

Query Match 27 5%; Score 440; DB 9; Length 447;
 Best Local Similarity 100.0%; Pred. No. 6e-200;
 Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1160 TGGAGCCCAAGGATGGAAGAGCTGTCTTTCAGACTGATGAGGAATGATGATGAGGAGG 1219
 Db 447 TGGAGCCCAAGGATGGAAGAGCTGTCTTTCAGACTGATGAGGAATGATGATGAGGAGG 388
 QY 1220 CGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1279
 Db 387 CGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 328
 QY 1280 CGGCCAGCTGGAAGGCCAGAGATGCCAGACCGGAGAGAGCCCTTATGACCCCTGAGGC 1339
 Db 327 CGGCCAGCTGGAAGGCCAGAGATGCCAGACCGGAGAGAGCCCTTATGACCCCTGAGGC 268
 QY 1340 GCGCTGTGCTTGT 1399
 Db 267 GCGCTGTGCTTGT 208
 QY 1400 GTCCAGAGCTCTGT 1459
 Db 207 GTCCAGAGCTCTGT 148
 QY 1460 TATGCTTTGGGGGAGACAGGCCAGAAATGAGTCTGTGTGTGTGTGTGTGTGTGTGTGT 1519
 Db 147 TATGCTTTGGGGGAGACAGGCCAGAAATGAGTCTGTGTGTGTGTGTGTGTGTGTGTGT 88
 QY 1520 CAGCTGCTTATGTAAAGGCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1579
 Db 87 CAGCTGCTTATGTAAAGGCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 28
 QY 1580 AAATGAAGTACATTCGCCAA 1599
 Db 27 AAATGAAGTACATTCGCCAA 8

RESULT 8
 AUI47968/c 587 bp mRNA linear EST 05-AUG-2002
 LOCUS AUI47968 MAMMAL Homo sapiens cDNA clone MAMMAL1002268 3', mRNA
 DEFINITION sequence.
 ACCESSION AUI47968
 VERSION AUI47968.1 GI:11009489
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 587)
 AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
 Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T., Sugano
 S., and Isogai,T.).
 HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S.,
 Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura
 Y., Nagai,T., Sugano,S., Isogai,T.)
 Unpublished (2000)

JOURNAL
 COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel.: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.
 Location/Qualifiers

FEATURES
 source
 1..587
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MAMMAL1002268"
 /clone_lib="MAMMAL1"
 /tissue_type="mammary gland"
 /note="Vector: pMT189FL3"

BASE COUNT 130 a 183 c 156 g 112 t 6 others

QY 1445 TGAGAGGTGAGGCTATGCTTTGGGGGACAGCCAGCAATGAACTGCTGGTCAGAGC 1504
 DB 167 TGAGAGGTGAGGCTATGCTTTGGGGGACAGCCAGCAATGAACTGCTGGTCAGAGC 108
 QY 1505 CCAGCTGGCTGGGGCCAGCTGCTATGTAAGGCTT 1540
 DB 107 CCAGCTGGCTGGGGCCAGCTGCTATGTAAGGCTT 72
 RESULT 12
 BM471470 416 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT 6478178 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563230
 DEFINITION 5' mRNA sequence.
 ACCESSION BM471470
 VERSION BM471470.1 GI:18520512
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 416)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM12293 row: 0 column: 07
 High quality sequence stop: 415.
 Location/Qualifiers
 1. 416
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5563230"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 BASE COUNT 95 a 118 c 125 g 78 t
 ORIGIN
 Query Match 25.0%; Score 399; DB 13; Length 416;
 Best Local Similarity 100.0%; Pred. No. 2.8e-180; Indels 0; Gaps 0;
 Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1200 GGAATTGATGCTTACGAGGCGGTCAGAGGCGCAGGACCCCAACTACTTCTGATGCT 1259
 DB 1 GGAATTGATGCTTACGAGGCGGTCAGAGGCGCAGGACCCCAACTACTTCTGATGCT 60
 QY 1260 CAGCGGTTGCTGAGACCCCGCCAGCTGGAAGCCCGCAGATGCCACCGCCAGAGA 1319
 DB 61 CAGCGGTTGCTGAGACCCCGCCAGCTGGAAGCCCGCAGATGCCACCGCCAGAGA 120
 QY 1320 GCCCTATGACCCCGGCGGCGGCTGCTAGTGTCTACTTGCAGGACCCCTGCTCT 1379
 DB 121 GCCCTATGACCCCGGCGGCGGCTGCTAGTGTCTACTTGCAGGACCCCTGCTCT 180
 QY 1380 TCCCTAGGAGCTCAGAGGCTGTCACAGCTCTGTGGGGGTGAGAGAGCTCTCTGGAG 1439
 DB 181 TCCCTAGGAGCTCAGAGGCTGTCACAGCTCTGTGGGGGTGAGAGAGCTCTCTGGAG 240
 QY 1440 AAGGTTGAGAAAGTGGAGGCTATGCTTTGGGGGACAGCCAGCAATGAACTGCTGGTCA 1499

DB 241 AAGGTTGAGAAAGTGGAGGCTATGCTTTGGGGGACAGCCAGCAATGAACTGCTGGTCA 300
 QY 1500 GGAGCCAGCTGCTGGGCGCCAGCTGCTATGTAAGGCTTCTAGTTCTGAGACC 1559
 DB 301 GGAGCCAGCTGCTGGGCGCCAGCTGCTATGTAAGGCTTCTAGTTCTGAGACC 360
 QY 1560 CCAGCCAGCAACCAATCAATTAAGTGAATGCCA 1598
 DB 361 CCAGCCAGCAACCAATCAATTAAGTGAATGCCA 399
 RESULT 13
 BE275818 603 bp mRNA linear EST 13-JUL-2000
 LOCUS 6011216161 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3346011 5',
 DEFINITION mRNA sequence.
 ACCESSION BE275818
 VERSION BE275818
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 603)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLM134 row: 0 column: 04
 High quality sequence stop: 599.
 Location/Qualifiers
 1. 603
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3346011"
 /clone_lib="NIH_MGC_20"
 /tissue_type="melanotic melanoma"
 /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 100 a 167 c 206 g 130 t
 ORIGIN
 Query Match 24.6%; Score 393; DB 10; Length 603;
 Best Local Similarity 99.6%; Pred. No. 2.2e-177; Indels 0; Gaps 0;
 Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 925 GTGTGCGCCGAGAGACTTGTGCTAGTCTGCGACTGCTGACCTGCGACCT 984
 DB 45 GTGTGCGCCGAGAGACTTGTGCTAGTCTGCGACTGCTGACCTGCGACCTGCGACCT 104
 QY 985 GAGATGTTTGTGCAACCCATGCGGCGCTGTCAGCTGCGCTGATGATCTGTTTACGTG 1044
 DB 105 GAGATGTTTGTGCAACCCATGCGGCGCTGTCAGCTGCGCTGATGATCTGTTTACGTG 164
 QY 1045 CGGGGCGAGTGTCTGCGCCATGCTGCTGCGGCGCTCTCCGCGCATGAGAGAGGCGAG 1104
 DB 165 CGGGGCGAGTGTCTGCGCCATGCTGCTGCGGCGCTCTCCGCGCATGAGAGAGGCGAG 224

Mon Jul 14 09:16:27 2003

us-09-784-810a-1_copy_2_1600.olg.rst

Page 10

OY	1105	CATVYGAGTAAATGACCCCTACTGTGGTTAATGTCGCCGGTGCGCTTCGCCTTGGAG	1165
Dd	225	CATATGAGATTAATAATGCCCCCTACTGTGTATGCTAACCTCCCGTAGCTCCGCTTGGAG	284
OY	1165	CCCAAGATGCGGAAGAGTGTGTTTTGACGTGATGCGGGAAATTGATGTTTACGAGGCGCGG	1224
Dd	289	CCCAAGATGCGGAAGAGTGTGTTTTGACGTGATGCGGGAAATTGATGTTTACGAGGCGCGG	344
OY	1225	CAGGCGCAGGTGACACCCAAACCTACTTGTGGATGAGTGAGCGGTGCGAGAGCCCGGCC	1284
Dd	345	CAGGCGCAGGTGACACCCAAACCTACTTGTGGATGAGTGAGCGGTGCGAGAGCCCGGCC	404
OY	1289	AGCTGGAAGCCCGCAGAGATGCGACCGCCGAGAAGAGCCCTTATGACCCCTGCGGCGCGCT	1344
Dd	405	AGCTGGAAGCCCGCAGAGATGCGACCGCCGAGAAGAGCCCTTATGACCCCTGCGGCGCGCT	464
OY	1345	GTCGCTTAGTGCTACTTGTGCAGAGACCCCTTCCTCCTCCCTAGAGGTGCGAGGCGCTGTCCA	1404
Dd	465	GTCGCTTAGTGCTACTTGTGCAGAGACCCCTTCCTCCTCCCTAGAGGTGCGAGGCGCTGTCCA	524
OY	1405	CAGCGCCCTGTCGGCGG 1449	
Dd	525	CAGCTCTCTGTGGGG 539	
RESULT 14	BMS50039	1115 bp mRNA linear EST 20-FEB-2002	
LOCUS	AGENCOURN_6544088 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745953		
DEFINITION	5', mRNA sequence.		
ACCESSION	BMS50039		
VERSION	BMS50039.1 GI:18789520		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgi.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapds-femail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LHAM12769 row: 1 column: 18 High quality sequence start: 20 High quality sequence stop: 451. Location/Qualifiers 1..1115 /organism="Homo sapiens" /db_xref="taxon:9606" /clone IMAGE:5745953" /clone_id="NIH MGC_118" /tissue_type="leukocyte" /lab_host="DH10B" /note="Vector: pCMW-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source: leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH-MGC Library."		
FEATURES	source		
BASE COUNT	177 a 361 c 385 g 192 t		
ORIGIN			

Query Match	Similarity	24.5%	Score 391	DB 13	Length 1115
Best local	Similarity	99.8%	Pred. No. 2.3e-176		
Matches	441	Conservative	0	Mismatches	1
				Indels	0
				Gaps	0
OY	18	AGCCACGGCTCCGGGGGGAAGGCGAGAGCCCCACAGCGCGCCCTGCGAGCGCCGCTGGG	77		
Db	203	AGCCACGGCTCCGGGGGGAAGGCGAGAGCCCCACAGCGCGCCCTGCGAGCGCCGCTGGG	262		
OY	78	CAGACCGCATTAAGAGAGCTGAAGGCGAGAGCGCGCGCACAGGGCACAGCCCCACAGCGC	137		
Db	263	CAGACCGCATTAAGAGAGCTGAAGGCGAGAGCGCGCGCACAGGGCACAGCCCCACAGCGC	322		
OY	138	AGGAGACCCCTGGCGACCGGAGAGCGCGGCTGAGAGTTATGATCTCAGCGGGGCCCCCG	197		
Db	322	AGGAGACCCCTGGCGACCGGAGAGCGCGGCTGAGAGTTATGATCTCAGCGGGGCCCCCG	382		
OY	198	GGGCGTGTCTCCCGCGGCGCCCTGTCGCGGTCGTGTGTGTGTGTGTGTGTGTGTGTGT	257		
Db	383	GGGCGTGTCTCCCGCGGCGCCCTGTCGCGGTCGTGTGTGTGTGTGTGTGTGTGTGTGT	442		
OY	258	CAGAGCCTTGACCTCTTCCGAGTCACTGTGACAGCCCTTTTGGTGTAGAGCTTAATATTC	317		
Db	443	CAGAGCCTTGACCTCTTCCGAGTCACTGTGACAGCCCTTTTGGTGTAGAGCTTAATATTC	502		
OY	318	CTTCAAGCTGATGCTCACTGAGAGCGGCGGAACCAAGCGCGGAGCTGTGTGTGTGTGTGT	377		
Db	503	CTTCAAGCTGATGCTCACTGAGAGCGGCGGAACCAAGCGCGGAGCTGTGTGTGTGTGTGT	562		
OY	378	GCTGGGCGCTGGGAGCGCTGT	437		
Db	563	GCTGGGCGCTGGGAGCGCTGT	622		
OY	438	GAACGGGCTCATGAGAGCGGCGCT	459		
Db	623	GAACGGGCTCATGAGAGCGGCGCT	644		

RESULT 15

BM808666

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM808666 1134 bp mRNA linear EST 05-MAR-2002

AGCNCOURT_5582812 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5471294

5' mRNA sequence.

BM808666

BM808666.1 GI:19125489

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 1134)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strassberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: DCTD/DPF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.lnl.gov>

Plate: L1CM1979 row: h column: 15

High quality sequence start: 18

High quality sequence stop: 559.

Location/Qualifiers

1..1134

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5471294"

/clone_lib="NIH_MGC_41"

/tissue_type="amelanotic melanoma, cell line"

/lab_host="PH10B (phage-resistant)"

FEATURES

source

/note="Organ: skin; Vector: POTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

BASE COUNT 186 a 382 c 373 g 193 t
ORIGIN

Query Match 24.5%; Score 391; DB 14; Length 1134;
Best Local Similarity 100.0%; Pred. No. 2.3e-176;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 18 AGCCACGGCTCCGGGCGGGGAAGAGCGAGCCGCCACAGCGGCCCTCGAGCGCCGCTGGG 77
    |||||||
Db 222 AGCCACGGCTCCGGGCGGGGAAGAGCGAGCCGCCACAGCGGCCCTCGAGCGCCGCTGGG 281
    |||||||
QY 78 CAGCACCGATAGAGAGCTGAAGGACGAGCGCGGCGGCGAGCGGCCGCCACAGCGGCC 137
    |||||||
Db 282 CAGCACCGATAGAGAGCTGAAGGACGAGCGCGGCGGCGAGCGGCCGCCACAGCGGCC 341
    |||||||
QY 138 AGGACCCCGCTGGCAGCGGAGCCCGGGTCTGAGGTATGATCCAGCGGGCGGCCCGC 197
    |||||||
Db 342 AGGACCCCGCTGGCAGCGGAGCCCGGGTCTGAGGTATGATCCAGCGGGCGGCCCGC 401
    |||||||
QY 198 GGGCGTGTCCCGCGGCCCTGCGGTGTGTGCTGTGAACCCGCGGCGGCAAGGG 257
    |||||||
Db 402 GGGCGTGTCCCGCGGCCCTGCGGTGTGTGCTGTGAACCCGCGGCGGCAAGGG 461
    |||||||
QY 258 CAAGGCTTTGACAGCTCTTCCGAGTCACTGAGCGCCCTTTGGCTGAGGCTGAATCTC 317
    |||||||
Db 462 CAAGGCTTTGACAGCTCTTCCGAGTCACTGAGCGCCCTTTGGCTGAGGCTGAATCTC 521
    |||||||
QY 318 CTTACAGCTGATGCTCACTGAGCGGCGGAAACCAACGCGGAGAGTGTGCGGTGAGAGA 377
    |||||||
Db 522 CTTACAGCTGATGCTCACTGAGCGGCGGAAACCAACGCGGAGAGTGTGCGGTGAGAGA 581
    |||||||
QY 378 GCTGGCGCGCTGGAGCGCTGTGTGTCATG 408
    |||||||
Db 582 GCTGGCGCGCTGGAGCGCTGTGTGTCATG 612
    |||||||
```

Search completed: July 12, 2003, 18:00:18
Job time : 2389 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 14:08:57 ; Search time 378 Seconds

(without alignments)
9526.310 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599
Sequence: 1 ccccccggggctctatagc.....aataaagtcattcccaa 1599

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N.Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1600	22	AAD14424
2	624	39.0	1438	22	AA159336
3	624	39.0	1533	20	AAV84490
4	624	39.0	1533	22	ABA83273
5	623	39.0	1562	24	ABL90618
6	623	39.0	1573	21	AAA75676
7	623	39.0	1821	22	AAH15652
8	590	36.9	1772	22	AAH16415
9	524	32.8	1719	22	AAH04477

10	452	28.3	1205	22	AAC84161	Human sphingosine
11	451	28.2	1173	22	ABL59532	Human sphingosine
12	440	27.5	587	24	AAH10443	Human CDNA clone (
13	420	26.3	1447	21	AAA50508	Human sphingosine
14	406	25.4	1200	22	AA161122	Human polynucleoti
15	381	23.8	675	22	AAH07214	Human CDNA clone (
16	288	18.0	604	22	AAH04943	Human CDNA clone (
17	240	15.0	240	22	AAD04478	Conserved region D
18	192	12.0	296	22	AA524477	Human ovarian PCR-
19	192	12.0	296	22	AAH83083	Human ovarian tumo
20	191	11.9	394	24	ABL66446	Lung cancer relate
21	60	3.9	60	24	ABN41630	Human SPHK1 CDNA c
22	30	1.9	30	24	ABL60764	Sphingosine kinase
23	25	1.6	25	21	AAA50512	Sphingosine kinase
24	25	1.6	25	21	AAA50513	Sphingosine kinase
25	25	1.6	1559	21	AA247168	Mouse sphingosine
26	25	1.6	1759	22	AAD14425	Mouse consensus sp
27	25	1.6	1815	21	AA247167	Mouse sphingosine
28	24	1.5	24	21	AAA50514	Sphingosine kinase
29	22	1.4	2932	24	ABK34670	Human CDNA for nov
30	21	1.3	27	24	ABL60763	Human SPHK1 CDNA c
31	21	1.3	4403765	22	AA196863	Mycobacterium tube
32	21	1.3	4411529	22	AA196862	Human SPHK1 CDNA c
33	20	1.3	20	24	ABL60762	Human S100A5 CDNA
34	20	1.3	333	21	AAAC81805	Human CAPL gene.
35	20	1.3	10952	17	AAV33345	Human CAPL gene.
36	20	1.3	10952	19	AAV41162	Human CAPL gene.
37	20	1.3	10952	20	AAK17696	Human CAPL genomic
38	19	1.2	26	22	AAAC84164	Human sphingosine
39	19	1.2	52	22	AAAC84164	Human sphingosine
40	19	1.2	278	14	AAO60903	Human brain Expres
41	19	1.2	363	23	AA570395	DNA encoding novel
42	19	1.2	363	23	AA570625	DNA encoding novel
43	19	1.2	363	23	AA578681	DNA encoding novel
44	19	1.2	363	23	AA592440	DNA encoding novel
45	19	1.2	382	23	AA566961	DNA encoding novel

ALIGNMENTS

RESULT 1	
AAD14424	
ID	AAD14424 standard; CDNA; 1600 BP.
AC	AAAD14424;
XX	
DT	01-NOV-2001 (first entry)
DE	Human sphingosine kinase (sphk) CDNA #1.
XX	
KW	Human; sphingosine kinase; Spk; restenosis; ischaemia; gene therapy;
KW	antisense therapy; cancer; sphingolipid; signalling molecule; apoptosis;
KW	cytostatic; tumour necrosis factor-alpha; TNF; atherosclerosis; lymphoma;
KW	leukaemia; vasotrophic; cell proliferative disorder; vascular disease; ss.
OS	Homo sapiens.
XX	
PH	Key
FT	Location/Qualifiers
FT	176..1330
FT	/*tag= a
FT	/product= "Human sphingosine kinase (sphk) protein #1"
XX	
PN	WO200160990-A2.
XX	
PD	23-AUG-2001.
XX	
PF	14-FEB-2001; 2001WO-US04789.
XX	
PR	14-FEB-2000; 2000US-0182360.
XX	
PR	22-MAR-2000; 2000US-0191261.
XX	
PA	(CURA-) CURAGEN CORP.

PA (GETH) GENENTECH INC.
XX
PI Rastelli L;
XX
DR WPI: 2001-514770/56.
DR P-PSDB: AAE07882.
XX
XX An isolated sphingosine kinase polypeptide useful for treating a
PT SphK-associated disorder especially cancer, restenosis or ischemia in a
PT human -
XX
PS Claim 8; Page 89-90; 107pp; English.
XX
XX The present invention relates to sphingosine Kinase (SphK) polypeptides
CC and nucleic acids encoding them. SphK is useful for treating a SphK-
CC associated disorder especially cancers such as leukaemia, lymphoma,
CC ovarian, breast, lung, colon, testicular, stomach and skin,
CC atherosclerosis, restenosis or ischaemia and cell proliferative disease
CC or disorder associated with vascular diseases. SphK gene is used in gene
CC therapy and antisense-therapy. Sphingolipids serving as signalling
CC molecules, have recently emerged as regulators of cell growth,
CC differentiation, diverse cell phenotypes and cell death. Activation of
CC SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
CC endothelial cells. The present sequence is human sphingosine kinase
CC (SphK) cDNA.
XX
XX Sequence 1600 BP; 265 A; 492 C; 531 G; 311 T; 1 other.

```
Query Match      100.0%; Score 1599; DB 22; Length 1600;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY	1	CCCCCGGGGGCTCTTATACCACAGSCTCCGGGGGGGAAGGGAGACCCCAACACGGAGCC	60
Db	2	CCCCCGGGGGCTCTTATACCACAGSCTCCGGGGGGGAAGGGAGACCCCAACACGGAGCC	61
OY	61	TGCGAGCGCCGCTGGGGGACGACGATATAGGAGCTGAAGGACGAGCGCGCACGGGC	120
Db	62	TGCGAGCGCCGCTGGGGGACGACGATATAGGAGCTGAAGGACGAGCGCGCACGGGC	121
OY	121	AAGCGCCCCACAGCGCCAGGAGACCCCTGGGACGGGAGCGCGGGGTGAGATTATGAT	180
Db	122	AAGCGCCCCACAGCGCCAGGAGACCCCTGGGAGGGAGACCGCGGGGTGAGATTATGAT	181
OY	181	CCAGCGGGAGCGCCCGGGGCGTGTCTCCGGCGCCCTGCGCGGTCTGTGCTGTGAAC	240
Db	182	CCAGCGGGAGCGCCCGGGGCGTGTCTCCGGCGCCCTGCGCGGTCTGTGCTGTGAAC	241
OY	241	CCGCGCGGCGGCAAGGGCAAGGCTTTGACGCTTTCGGAGTCACTGACGTCACGCCCTTTG	300
Db	242	CCGCGCGGCGGCAAGGGCAAGGCTTTGACGCTTTCGGAGTCACTGACGTCACGCCCTTTG	301
OY	301	GCTGAGGCTGAAATTCCTTCACGCTGATAGTCACTGAGCGCGGAGACAGAGCGGGAG	360
Db	302	GCTGAGGCTGAAATTCCTTCACGCTGATAGTCACTGAGCGCGGAGACAGAGCGGGAG	361
OY	361	CTGTGCGGTGGAGAGAGCTGGGCGCGCTGGAGACCTGTGTGTATGTGTGAGAGCGG	420
Db	362	CTGTGCGGTGGAGAGAGCTGGGCGCGCTGGAGACCTGTGTGTATGTGTGAGAGCGG	421
OY	421	CTGATGACAGAGTGTGTGAACGGGCTCATATGAGAGCGGCTGACATGTGGAGACCGGCATTCAG	480
Db	422	CTGATGACAGAGTGTGTGAACGGGCTCATATGAGAGCGGCTGACATGTGGAGACCGGCATTCAG	481
OY	481	AAGCCCCGTGTAGCTCCCGACAGAGCTGTGGCAAGCGGCTGGGAGCTTCTCTTGAAACAT	540
Db	482	AAGCCCCGTGTAGCTCCCGACAGAGCTGTGGCAAGCGGCTGGGAGCTTCTCTTGAAACAT	541
OY	541	TATGTGTGGCTATGAGCAGGTCAACCAATAGAGACTCTCTGACCACTGCAAGCTATTTGCTG	600
Db	542	TATGTGTGGCTATGAGCAGGTCAACCAATAGAGACTCTCTGACCACTGCAAGCTATTTGCTG	601
OY	601	TGCGCCCGGGTGTCTACCCATGAACCTGTGTCTCTGACAGGGCTTTCGGGGCTGGCG	660

Db	602	TECCGCCCGGTGCTGCACCCATGACCTCTCTCTCTGACACAGCCCTTGCGGGCGTGGCC	661
Oy	661	TGCTTCTCTGTGCTCAGCCCTGGACCTGGGGGCTTCAATTCGATGTGACACTAGAGATAT	720
Db	662	TGCTTCTGTGTCTCAGCTGACCTGGGGGCTTCAATTCGATGTGACACTAGAGATAT	721
Oy	721	AAGTATGGGGGTGTGGGGGAGATGGGTACTCTGGGACCTTCTGCTGGAGCC	780
Db	722	AAGTATGGGGGTGTGGGGGAGATGGGTACTCTGGGACCTTCTGCTGGAGCC	781
Oy	781	CTGGGCACCTAACCGCGGCGTATGGTCTACCTCCCTGTAGAGAAAGTGGTTCAAGACA	840
Db	782	CTGGGCACCTAACCGCGGCGTATGGTCTACCTCCCTGTAGAGAAAGTGGTTCAAGACA	841
Oy	841	CCCTGCTTCCCGGTGTGGTCCAGTAGGGCCCGGTATGACACACTGTGTCCACTGGAG	900
Db	842	CCCTGCTTCCCGGTGTGGTCCAGTAGGGCCCGGTATGACACACTGTGTCCACTGGAG	901
Oy	901	GAGCAGGTGCTTCTCTACGTGGAGGTGGTCCCGCAGAGAGACTTGTGTAGTCCCTGGCA	960
Db	902	GAGCAGGTGCTTCTCTACGTGGAGGTGGTCCCGCAGAGACTTGTGTAGTCCCTGGCA	961
Oy	961	CTGTGCACATCGACCTGGGCGCAGATATTTGTGTGCAACCAATAGGGGCGGTGTAGCT	1020
Db	962	CTGTGCACATCGACCTGGGCGCAGATATTTGTGTGCAACCAATAGGGGCGGTGTAGCT	1021
Oy	1021	GGCGTCATGCATCTGTCTACGTGGGCGGAGGTCTCTGTCCATGTCTGTGGCGCTC	1080
Db	1022	GGCGTCATGCATCTGTCTACGTGGGCGGAGGTCTCTGTCCATGTCTGTGGCGCTC	1081
Oy	1081	TTTCTGGGCACTGAGAGAGGGCAGGCATATGAGTATGATGGTCCCTACTTGTATATGTG	1140
Db	1082	TTTCTGGGCACTGAGAGAGGGCAGGCATATGAGTATGATGGTCCCTACTTGTATATGTG	1141
Oy	1141	CCCGTGGCGCTCTCCGCTTGTGAGGCCCAAGATGGGAAAGTGTGTGACATGATGGG	1200
Db	1142	CCCGTGGCGCTCTCCGCTTGTGAGGCCCAAGATGGGAAAGTGTGTGACATGATGGG	1201
Oy	1201	GAAATTGATGGTTAGGAGGCGGTGAGGGCCAGGTGACACCAAACTACTTCTGATATGTC	1260
Db	1202	GAAATTGATGGTTAGGAGGCGGTGAGGGCCAGGTGACACCAAACTACTTCTGATATGTC	1261
Oy	1261	AACGGTTGCTGTGAGGCCCGCCGACGTGTGAAACCCCAAGATGCCACGCCAGAAAGG	1320
Db	1262	AACGGTTGCTGTGAGGCCCGCCGACGTGTGAAACCCCAAGATGCCACGCCAGAAAGG	1321
Oy	1321	CCCTTATATACCCCTGTGGGACCGCTGTGCTTATGTCTACTTTCAGAGACCCCTTCTCTT	1380
Db	1322	CCCTTATATACCCCTGTGGGACCGCTGTGCTTATGTCTACTTTCAGAGACCCCTTCTCTT	1381
Oy	1381	CCCTTGGGCTTGCAGAGGCTGTGCACAGCTCCTTGGGGGTGAGAGACATCTCTTGAGA	1440
Db	1382	CCCTTGGGCTTGCAGAGGCTGTGCACAGCTCCTTGGGGGTGAGAGACATCTCTTGAGA	1441
Oy	1441	AAGGTGAGAAAGGTGAGGCTATGTCTTGGGGGACAGGCGCAATGAAGATCGGGTGG	1500
Db	1442	AAGGTGAGAAAGGTGAGGCTATGTCTTGGGGGACAGGCGCAATGAAGATCGGGTGG	1501
Oy	1501	GAGCCACAGCTGGTGGGACCACAGCTGCTATGTAAAGGCTTCTAGTTGTCTGAGACCC	1560
Db	1502	GAGCCACAGCTGGTGGGACCACAGCTGCTATGTAAAGGCTTCTAGTTGTCTGAGACCC	1561
Oy	1561	CACCCACAGCAACAATTCAAATTAAGTGAATTTCCCA	1599
Db	1562	CACCCACAGCAACAATTCAAATTAAGTGAATTTCCCA	1600

```

RESULT 2
AA159336
ID AA159336 standard; CDNA; 1438 BP
XX
AC AA159336;

```


P1	Fan P, Feng P, Ferlie AM, Fischer CL, Florence C;
P1	Florence K, Greene JM, Hu J, Kyaw H, Lefleur DW;
P1	Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
P1	Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
XX	
DR	NP1: 1999-059865/05.
DR	p-FSDB; AAM86613.
KA	New isolated human genes and the secreted polypeptides they encode -
FT	useful for diagnosis and treatment of e.g. cancers, neurological
P7	disorders, immune diseases, inflammation or blood disorders
PS	
XX	Claim 4, Page 343; 772pp; English.
CC	The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
CC	encoding human secreted proteins (AAM85534 to AAM88756). The secreted
CC	protein gene sequences are deposited with the ATCC under deposit numbers
CC	ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC	209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC	cells comprising recombinant vectors containing the nucleic acid
CC	sequences are used for the recombinant production of the secreted
CC	proteins. The polynucleotide and amino acid sequences are useful for are
CC	useful for preventing, treating or ameliorating medical conditions e.g.
CC	by protein or gene therapy. Pathological conditions can be also
CC	diagnosed by determining the amount of the new polypeptides in a sample
CC	or by determining the presence of mutations in the new polynucleotides.
CC	Specific uses are described for each of the polynucleotides, based on
CC	which tissues they are most highly expressed in, and include developing
CC	products for the diagnosis or treatment of cancer, neurodegenerative
CC	disorders, developmental abnormalities and fetal deficiencies, blood
CC	diseases, tumours, leukemias, diseases of the immune system, autoimmune
CC	diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC	ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC	restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC	such as osteoporosis, arthritis or malignancies, infections and AIDS. The
CC	lung or thymus, digestive/endocrine disorders, diseases of testes.
CC	The polypeptides are also useful for identifying their binding partners.
CC	The present sequence represents a gene encoding a human secreted protein
CC	(see descriptor line for gene number and clone identification).
XX	
SQ	Sequence 1533 BP: 272 A; 454 C; 495 G; 302 T; 10 other:
Query Match	39.0%; Score 624; DB 20; Length 1533;
Best Local Similarity	99.9%; Pred. No. 7.6e-772;
Matches 674; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
DY	925 GTGTGTCGCCAGCAGACCTTTGTGTACTGTCGGCACTCTGCACCTCGCCGCACT 984
Db	824 GTTGGTGCCCGACAGACACTTTGTGTACTGTCGGCACTCTGCACCTCGCCGCACT 883
OY	985 GAGATGTTCTCTCACCCATGCGGCGCGTGGACAGTGGCATGTCATTGTTCTACGTG 1044
Db	884 GAGATGTTCTCTCACCCATGCGGCGCGTGGACAGTGGCATGTCATTGTTCTACGTG 943
OY	1045 CGCGCGGGAATGTCCTGTGTCGTCATGTCGTGCGCTTCCTCGGCATGGAAGGGGCAAG 1104
Db	944 CGCGCGGGAATGTCCTGTGTCGTCATGTCGTGCGCTTCCTCGGCATGGAAGGGGCAAG 1003
OY	1105 CATATGAGATTAATATGCCCCCTACTTGGATATATGACCCTGTCGCTTCGCTTGAG 1164
Db	1004 CATATGAGATTAATTAATGCCCCCTACTTGGATATATGACCCTGTCGCTTCGCTTGAG 1063
OY	1165 CCCAAGAGATGGGAAAAGTGTGTTGGCAAGTGGAGGGGAATTGATGTTTAGCGAGGCCGTG 1224
Db	1064 CCCAAGAGATGGGAAAAGTGTGTTGGCAAGTGGAGGGGAATTGATGTTTAGCGAGGCCGTG 1122
OY	1225 CAGGGCCAGGTGAGCCCAAACACTGTCGGATGGTGTAGCGGTGGAGGCCGCCGCC 1284
Db	1124 CAGGGCCAGGTGAGCCCAAACACTGTCGGATGGTGTAGCGGTGGAGGCCGCCGCC 1188
OY	1285 ACCTGGAAGCCCACGACATGTCACCGCCAGAAGAACCTTATATACCCCTGGGCGCGCT 1344
Db	1184 ACCTGGAAGCCCACGACATGTCACCGCCAGAAGAACCTTATATACCCCTGGGCGCGCT 1243

DB 1484 AAGTGACATGCCAA 1498

RESULT 5

AB190618/c
ID ABL90618 standard; cDNA; 1562 BP.

XX ABL90618:

XX 24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 1180.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX antiviral; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiast; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; gene; ss.

XX Homo sapiens.

XX WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US16450.

XX 19-MAY-2000; 2000US-205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

XX P-PSDB; ABB90209.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
XX prevention of neural, immune system, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX disorders.

XX Claim 4; SEQ ID NO 1180; 2081bp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (AB89040-AB89044) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WPI at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 1562 BP; 328 A; 499 C; 459 G; 273 T; 3 other:

XX Query Match 39.0%; Score 623; DB 24; Length 1562;

XX Best Local Similarity 99.9%; Pred. No. 2.1e-271;

XX Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 925 GTGGTGGCCGACGAGACTTTGTGCTAGCTGCGACCTGCGACCTGGGCGAGT 984

OY 712 GTGGTGGCCGACGAGACTTTGTGCTAGCTGCGACCTGCGACCTGGGCGAGT 653

OY 985 GAGATGTTTCTGCAACCCATGGGCGCTGTGACGCTGGCTGATCATCTTACGCTG 1044

DB 652 GAGATGTTTCTGCAACCCATGGGCGCTGTGACGCTGGCTGATCATCTTACGCTG 593

OY 1045 CGGGCGGAGAGTGTCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104

DB 592 CGGGCGGAGAGTGTCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 533

OY 1105 CATATGAGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1164

DB 532 CATATGAGATGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 473

OY 1165 CCCAAGATGAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1224

DB 472 CCCAAGATGAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 413

OY 1225 CAGGCGCAGGTGACCCCAACTACTTCTGATGCTGACGCTGCTGACGCTGCTGACGCTGCTG 1284

DB 412 CAGGCGCAGGTGACCCCAACTACTTCTGATGCTGACGCTGCTGACGCTGCTGACGCTGCTG 353

OY 1285 AGCTGGAAGCCCGACGAGATGCTGACGAGATGCTGACGAGATGCTGACGAGATGCTGACG 1344

DB 352 AGCTGGAAGCCCGACGAGATGCTGACGAGATGCTGACGAGATGCTGACGAGATGCTGACG 293

OY 1345 GTGCTTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1404

DB 292 GTGCTTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 233

OY 1405 CAGCTCCTGTGGGGGTGGAGAGTCTCTGAGAGAGGATGAGAGGATGAGAGGATGAGAGG 1464

DB 232 CAGCTCCTGTGGGGGTGGAGAGTCTCTGAGAGAGGATGAGAGGATGAGAGGATGAGAGG 173

OY 1465 TTTGGGGGAGCAGGAGATGAGTCTGAGAGAGGATGAGAGGATGAGAGGATGAGAGGATG 1524

DB 172 TTTGGGGGAGCAGGAGATGAGTCTGAGAGAGGATGAGAGGATGAGAGGATGAGAGGATG 113

OY 1525 GCGTATGATGAGGCTCTGCTGCTGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATG 1584

DB 112 GCGTATGATGAGGCTCTGCTGCTGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATG 53

OY 1585 AAGTGACATGCCCA 1598

DB 52 AAGTGACATGCCCA 39

RESULT 6

AAA75676
ID AAA75676 standard; DNA; 1573 BP.

XX AAA75676:

XX 22-JAN-2001 (first entry)

XX DNA encoding a human regulator of intracellular phosphorylation.

XX Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;

XX neurological disorder; Parkinson's disease; demyelinating disease;

XX meningitis; developmental disorder; neuromuscular disorder; cancer;

XX myasthenia gravis; cell proliferative disorder; actinic keratosis;

XX arteriosclerosis; atherosclerosis; leukemia; melanoma; bronchitis;

XX autoimmune disorder; inflammatory disorder; Addison's disease;

XX acquired immunodeficiency disease; allergy; diabetes mellitus;

XX rheumatoid arthritis; microbial infection; trauma; ss.

XX Homo sapiens.

XX WO200055332-A2.

XX 21-SEP-2000.

Key Location/Qualifiers
CDS 130..1284
/tag- a
/product- "regulator of intracellular phosphorylation"

```

XX 17-MAR-2000; 2000WO-US07277.
XX
XX 18-MAR-1999; 99US-0125593.
XX 20-MAY-1999; 99US-0135049.
XX 09-JUL-1999; 99US-0143188.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Tang YF, Yue H, Hillman JL, Baughn MR, Azimzai Y,
XX Lu DM, Au-Young J;
XX WPI: 2000-602121/57.
XX P-PSDB; AAB18659.
XX
XX Novel human intracellular phosphorylation regulator polypeptides and
XX polynucleotides for diagnosis, prevention and treatment of
XX neurological, cell proliferative and autoimmune/inflammatory disorders
XX
XX
XX PS Claim 4; Page 89; 96pp; English.
XX
XX The present sequence encodes a human regulator of intracellular
XX phosphorylation (HRIP). HRIP is useful for screening agonists and
XX antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
XX are useful for treating a disease or condition associated with
XX decreased or increased expression of functional HRIP. Diseases treated
XX or diagnosed include neurological disorders such as stroke, Parkinson's
XX disease, demyelinating diseases, bacterial and viral meningitis and
XX other developmental disorders of the central nervous system,
XX neuromuscular disorders, myasthenia gravis, cell proliferative disorders
XX such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
XX including leukemia, melanoma, myeloma and cancer of the adrenal gland,
XX bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
XX inflammatory disorder such as Addison's disease, acquired
XX immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
XX rheumatoid arthritis, microbial infection and trauma.
XX
XX Sequence 1573 BP; 274 A; 480 C; 514 G; 305 T; 0 other:
XX
XX
XX Query Match 39.0%; Score 623; DB 21; Length 1573;
XX Best Local Similarity 99.9%; Pred. No. 2,1e-271;
XX Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 925 GTGTCGCCGAGGAGGAGCTTGTCTAGTCTCTGACACTGCGACCTGCGACCTGGCAGT 984
XX |||||||
XX 880 GTGTCGCCGAGGAGGAGCTTGTCTAGTCTCTGACACTGCGACCTGCGACCTGGCAGT 939
XX |||||||
XX 985 GAGATGTTTGTCTGACCATGAGGCGCTGTGACAGTGGCGTCATGATCTGTTACGTG 1044
XX |||||||
XX 940 GAGATGTTTGTCTGACCATGAGGCGCTGTGACAGTGGCGTCATGATCTGTTACGTG 999
XX |||||||
XX 1045 CGGCGGAGAGTGTCTGTCATGCTGCGCCTCTTCTGCGCATGAGAGGCGCAG 1104
XX |||||||
XX 1000 CGGCGGAGAGTGTCTGTCATGCTGCGCCTCTTCTGCGCATGAGAGGCGCAG 1059
XX |||||||
XX 1105 CATATGAGATGATGATGCCCCCTACTGTTATATGTCGCCGCGTGCGCTTCCGTTGGAG 1164
XX |||||||
XX 1060 CATATGAGATGATGATGCCCCCTACTGTTATATGTCGCCGCGTGCGCTTCCGTTGGAG 1119
XX |||||||
XX 1165 CCCAAGATGAGGAGGAGTGTCTGACAGTGGAGGGAATGATGAGTGAAGGAGCCCTG 1224
XX |||||||
XX 1120 CCCAAGATGAGGAGGAGTGTCTGACAGTGGAGGGAATGATGAGTGAAGGAGCCCTG 1179
XX |||||||
XX 1225 CAGGCGCAGTGTGACCCCAACTCTTGTGATGTGACAGGTTGCGTGGAGCCCGCC 1284
XX |||||||
XX 1180 CAGGCGCAGTGTGACCCCAACTCTTGTGATGTGACAGGTTGCGTGGAGCCCGCC 1239
XX |||||||
XX 1285 AGCTGGAAGCCCGACGAGATGCCACCGCAGAGAGCCCTATGACCCCTGGGCGGCT 1344
XX |||||||
XX 1240 AGCTGGAAGCCCGACGAGATGCCACCGCAGAGAGCCCTATGACCCCTGGGCGGCT 1299
XX |||||||
XX 1345 GTGCTTAGTGTCTACTTGACAGACCCCTTCCCTCCCTAGGAGCTGACAGGCGCTGTGCA 1404

```

```

DB 1300 GTGCTTAGTGTCTACTTGCAGAGACCTTCTCTCTTCCCTAGAGGCTGACAGGCGCTGTGCA 1359
DB 1405 CAGTCTCTGTGGGGGTGAGAGACCTCTCTGAGAGAGGTTGAGAACAGTGGAGCTATGC 1464
DB 1360 CAGTCTCTGTGGGGGTGAGAGACCTCTCTGAGAGAGGTTGAGAACAGTGGAGCTATGC 1419
DB 1465 TTGCGGGGACAGGCGCAAGTGAAGTGGTGGTGGAGAGCCACAGCTGGTGGGCGCAGCT 1524
DB 1420 TTGCGGGGACAGGCGCAAGTGAAGTGGTGGTGGAGAGCCACAGCTGGTGGGCGCAGCT 1479
DB 1525 GCTATGTGAAGCCCTTCTAGTGTGAGAGCCCGCAGCCGACGAGAACAAATCAATA 1584
DB 1480 GCTATGTGAAGCCCTTCTAGTGTGAGAGCCCGCAGCCGACGAGAACAAATCAATA 1539
DB 1585 AAGTGACATTCCTCA 1598
DB 1540 AAGTGACATTCCTCA 1553

RESULT 7
AAH15652
ID AAH15652 standard; CDNA; 1821 BP.
AC AAH15652;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA sequence SEQ ID NO:13996;
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 13996; 2537bp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification, where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and

```


Db 1560 TTGGGGGGAGAGCCAGATGAGTCTGGGTACAGAGCCAGCTGGGCCCCAGCT 1619

Qy 1525 GCCTATGTAAGCCCTTCTAGTTTGT 1550
 1520 GCCTATGTAAGCCCTTCTAGTTTGT 1645

Db 1620 GCCTATGTAAGCCCTTCTAGTTTGT 1645

RESULT 10
 AAC84161
 ID: AAC84161 standard; cDNA: 1205 BP.

XX AAC84161;
 AC AAC84161;
 XX 19-MAR-2001 (first entry)
 DT Human sphingosine kinase encoding cDNA.
 XX
 DE Sphingosine kinase; antiarthritic; antiasthmatic; antiarteriosclerotic;
 XX antiinflammatory; neuroprotective; antibacterial; immunosuppressive;
 KW human; ss.
 KM
 XX Homo sapiens.
 OS

PH Key Location/Qualifiers
 FT CDS 33..1187
 FT /tag="a
 FT /product="sphingosine kinase"

XX MO200070028-A1.
 XX 23-NOV-2000.
 XX
 XX 12-MAY-2000; 2000MO-AU00457.
 XX
 XX 13-MAY-1999; 99AU-0000339.
 XX
 XX 08-JUL-1999; 99AU-0001504.
 XX
 XX (JOHJ) JOHNSON & JOHNSON RES PTY LTD.
 XX
 XX Pilsen SM, Wattenberg BW, Xia P, D'Andrea RJ, Gamble JR, Vadas MA;
 XX
 XX WPI: 2001-016227/02.
 XX
 XX P-PSDB: AAB48007.

DR Novel sphingosine kinase protein and nucleic acid molecules for
 XX diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,
 XX atherosclerosis, inflammation, meningitis, multiple sclerosis and
 XX septic shock
 PT
 XX
 XX Claim 4: Fig 7a; 100pp; English.

XX This cDNA encodes a human sphingosine kinase (SK) protein. The human SK
 XX protein, encoding nucleic acids and modulators are useful for modulating
 XX expression, functional activity or cellular functional activity of
 XX sphingosine kinase in a subject and also for treating a mammal by
 XX modulating the activity of SK. Diseases treated by regulating SK
 XX cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
 XX inflammation, meningitis, multiple sclerosis and septic shock.
 XX
 XX Sequence 1205 BP; 196 A; 357 C; 400 G; 252 T; 0 other;

SO Query Match 28.3%; Score 452; DB 22; Length 1205;
 Best Local Similarity 100.0%; Pred. No. 3.7e-194; Indels 0; Gaps 0;
 Matches 452; Conservative 0; Mismatches 0;

Qy 156 GAGAGCCGGGTGAGATTATGATCCAGCGGCGGCCCGGGGCTGCTCCGGCGGCC 215
 156 GAGAGCCGGGTGAGATTATGATCCAGCGGCGGCCCGGGGCTGCTCCGGCGGCC 215
 14 GAGAGCCGGGTGAGATTATGATCCAGCGGCGGCCCGGGGCTGCTCCGGCGGCC 73
 14 GAGAGCCGGGTGAGATTATGATCCAGCGGCGGCCCGGGGCTGCTCCGGCGGCC 73
 Qy 216 CTGCGCGGTGCTGCTGCTGAACCCGCGCGCGGCAAGGCGGCTTGCACCTCTT 275
 216 CTGCGCGGTGCTGCTGCTGAACCCGCGCGCGGCAAGGCGGCTTGCACCTCTT 275
 74 CTGCGCGGTGCTGCTGCTGAACCCGCGCGCGGCAAGGCGGCTTGCACCTCTT 133

Qy 276 CCGAGTACAGTGCAGACCCCTTTGCTGAGGCTGAAATCTCTTCAACGCTGATGCTCAC 335
 276 CCGAGTACAGTGCAGACCCCTTTGCTGAGGCTGAAATCTCTTCAACGCTGATGCTCAC 335
 Db 134 CCGAGTACAGTGCAGACCCCTTTGCTGAGGCTGAAATCTCTTCAACGCTGATGCTCAC 193
 134 CCGAGTACAGTGCAGACCCCTTTGCTGAGGCTGAAATCTCTTCAACGCTGATGCTCAC 193

Qy 336 TGAGGGGGGGAACACACGCGCGGAGCTGTGCGGTGCGAGACCTGGCGCGGAGAGC 395
 336 TGAGGGGGGGAACACACGCGCGGAGCTGTGCGGTGCGAGACCTGGCGCGGAGAGC 395
 Db 194 TGAGGGGGGGAACACACGCGCGGAGCTGTGCGGTGCGAGACCTGGCGCGGAGAGC 253
 194 TGAGGGGGGGAACACACGCGCGGAGCTGTGCGGTGCGAGACCTGGCGCGGAGAGC 253

Qy 396 TCTGTGTGATGCTGTGAGACGGGCTGATGACAGAGGTGTAACGGGCTCATGAGAGG 455
 396 TCTGTGTGATGCTGTGAGACGGGCTGATGACAGAGGTGTAACGGGCTCATGAGAGG 455
 Db 254 TCTGTGTGATGCTGTGAGACGGGCTGATGACAGAGGTGTAACGGGCTCATGAGAGG 313
 254 TCTGTGTGATGCTGTGAGACGGGCTGATGACAGAGGTGTAACGGGCTCATGAGAGG 313

Qy 456 GCCTGACTGGAGACCGCCATCCAGACGCCCTGTGTAGCTTCCACAGAGGCTTGACAA 515
 456 GCCTGACTGGAGACCGCCATCCAGACGCCCTGTGTAGCTTCCACAGAGGCTTGACAA 515
 Db 314 GCCTGACTGGAGACCGCCATCCAGACGCCCTGTGTAGCTTCCACAGAGGCTTGACAA 373
 314 GCCTGACTGGAGACCGCCATCCAGACGCCCTGTGTAGCTTCCACAGAGGCTTGACAA 373

Qy 516 CGGCTGTGAGAGCTTCTTGAACCAATTATGCTGATGAGAGGTACCAATGAAGACT 575
 516 CGGCTGTGAGAGCTTCTTGAACCAATTATGCTGATGAGAGGTACCAATGAAGACT 575
 Db 374 CGGCTGTGAGAGCTTCTTGAACCAATTATGCTGATGAGAGGTACCAATGAAGACT 433
 374 CGGCTGTGAGAGCTTCTTGAACCAATTATGCTGATGAGAGGTACCAATGAAGACT 433

Qy 576 CCGTACCACACTGCACGCTATTGCTGCGGCC 607
 576 CCGTACCACACTGCACGCTATTGCTGCGGCC 607
 Db 434 CCGTACCACACTGCACGCTATTGCTGCGGCC 465

RESULT 11
 ABL59532
 ID: ABL59532 standard; cDNA: 1173 BP.

XX ABL59532;
 AC ABL59532;
 XX 16-JUL-2002 (first entry)
 DT Human sphingosine kinase (SphK1) cDNA SEQ ID NO:32.
 XX
 XX Human sphingosine kinase; SphK1; enzyme; chromosome 17q25.2; gene;
 XX tumour; lipid associated gene; lipid metabolism; lipid synthesis; ss.
 KW
 KM
 XX Homo sapiens.
 OS
 XX MO200227028-A1.
 XX
 XX 04-APR-2002;
 XX
 XX 27-SEP-2001; 2001MO-US30366.
 XX
 XX 28-SEP-2000; 2000US-0676052.
 XX
 XX (ATAT-) ATAIRGIN TECHNOLOGIES INC.
 XX
 XX Skinner MK, Patton JL, Chaudhary J;
 XX
 XX WPI: 2002-402054/43.

DR Identifying tumor characteristics in a tissue sample taken from a
 XX patient, involves determining the copy number or expression level of
 XX genes associated with lipid metabolism, synthesis or action
 XX
 XX Example 1; Page 90; 113pp; English.

XX The present invention describes a method for identifying tumour
 XX characteristics, comprising measuring a copy number or expression level
 XX of at least two genes associated with lipid metabolism, synthesis, or
 XX action in cells from a patient tissue sample, and comparing the results
 XX with a copy number or expression level of the genes in a normal cell.
 XX Also described is an array of nucleic acid polymers immobilised on a
 XX solid support, comprising a solid support, at least two different nucleic
 XX acid polymers which are each specific for a different gene associated
 XX with lipid metabolism, synthesis or action, where each nucleic acid
 XX polymer is located at a predetermined position on the solid support, and
 XX the array comprises nucleic acid polymers which are specific for less
 XX than 100 genes other than the selected genes. The method is useful for
 XX determining tumour characteristics in a tissue sample taken from a

PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0653036.
PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPT; 2001-442253/47.
DR P-PSDB; AAM41966.

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

PS Claim 1; SEQ ID NO 5111; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM86642-AAM42213) with neurotrophic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

SQ Sequence 1200 BP; 218 A; 353 C; 375 G; 254 T; 0 other;

Query Match	25.48;	Score 406;	DB 22;	Length 1200;
Best Local Similarity	99.68;	Pred. No. 2.2e-173;		

QY	925	GTGGTGGCCGACGAGGACTTTGTCTATGCTCTGGACATCGGACATCGACACTGGCCAGT	984
QY	516	GTGGTGGCCGACGAGGACTTTGTCTATGCTCTGGACATCGGACATCGACACTGGCCAGT	575
Db	985	GAGATGTTTCTGCAACCCATGAGGCGCGGTGCTGACCTGGCGCATGCTGTTCTAACGTG	1044
QY	576	GAGATGTTTCTGCAACCCATGAGGCGCGGTGCTGACCTGGCGCATGCTGTTCTAACGTG	635
Db	1045	CGGGCGGGAGTGTCTGTCGCCATGCTGCTGCGCCTTTCCTGCGCATGAGAAAGGCGAG	1104
QY	636	CGGGCGGGAGTGTCTGTCGCCATGCTGCTGCGCCTTTCCTGCGCATGAGAAAGGCGAG	695
Db	1105	CATATGGAGTATGAAATGCCCTACTTGTGATATATGCCCCGTCGTGCTTCCGCTTGGAG	1166
QY	696	CATATGGAGTATGAAATGCCCTACTTGTGATATATGCCCCGTCGTGCTTCCGCTTGGAG	755
Db	1165	CCCAAGAGATGGAAAGGTTGTTGCAATGGATGGGGAATTGATGGTTAAGGAGGCCGTG	1224
QY	756	CCCAAGAGATGGAAAGGTTGTTGCAATGGATGGGGAATTGATGGTTAAGGAGGCCGTG	815
Db	1225	CAGGCGCAGGTGACCCCAACTACTTTCGTGATGTCACGCGGTGGCTGGAGACCCCGCC	1284
QY	816	CAGGCGCAGGTGACCCCAACTACTTTCGTGATGTCACGCGGTGGCTGGAGACCCCGCC	875
Db	1285	AGCTGGAAGCCCCAGCAGATGCCACCGCCAGAAAGCCCTTATGACCCCTGGGCGGCGCT	1344
QY	876	AGCTGGAAGCCCCAGCAGATGCCACCGCCAGAAAGCCCTTATGACCCCTGGGCGGCGCT	935
Db	1345	GTGCTTAAGTGTCTACTTTCGAGGACCTTCTCTCTTCCCTAAGGGTGCAGGGCTGTGCA	1404

Db 936 GTGCCTTAGTGTCTACTGTCAGGACCCCTTCCTCCCTAAGGCTGAGGGCTGTCCA 995

QY 1405 CAGCTCCCTGGGGGGTGGAGGAGACTCC 1432

Dd 996 CAGCTCCCTGGGGGGTGGAGGAGACTCC 1033

RESULT 15

ID AAH07214 standard; cDNA; 675 BP

AC AAH07214;

DT 26-JUN-2001 (first entry)

Human cdna clone (5'-primer) SEQ ID NO:4049.

	Human; primer
KW	
XY	

OS Homo sapiens.
XX

PN EP10/401/-AZ
XX

FD 07 FEB 2001.
XX

[illegible]

PR 27-AUG-1999; 99JP-0300253.

PR 02-MAY-2000; 2000JP-0183767

XX
XX

XX
DT
XX
C
E
t
e
t
E
W
a
b
t
l
c
r
r

PI Ishii S, Sugiyama T,
YV

DR WPI; 2001-318749/34.
XX

PT	Primer sets for synthesis of full-length cDNAs
PT	

PT full-length cDNAs - and/or diagnosis of the abnormality of the proteins encoded by the

PS Claim 1; SEQ ID 4049; 2537pp + CD ROM; English..

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to
CC AA895893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

Sequence 675 BP; 100 A; 234 C; 240 G; 98 T; 3 other;
Query Match 23.8%; Score 381; DB 22; Length 675;

Best Local Similarity 100.0%; Pred. No. 4.6e-162;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	18	AGCCACGGGCTCCGGGGGAGAGGAGCCACAGCCGCTGCGAGCCCGGCTGAG	77
Db	193	AGCCACGGGCTCCGGGGGAGAGGAGCCACAGCCGCTGCGAGCCCGGCTGAG	252
OY	78	CAGCAGCATTAAGAGCTGAAGCAGAGCCCGCCACAGCGCCCGCCACAGCGCC	137
Db	253	CAGCAGCATTAAGAGCTGAAGCAGAGCCCGCCACAGCGCCCGCCACAGCGCC	312
OY	138	AGGACCCCGCTGGAGAGGAGCGCGGGGTGAGGTTATGATCCAGCGGGCGCCCG	197
Db	313	AGGACCCCGCTGGAGAGGAGCGCGGGGTGAGGTTATGATCCAGCGGGCGCCCG	372
OY	198	GGGCGTGTCCCGGCGCTGCGCTGCTGTGCTGAACCGCGGGCGGCAAGGG	257
Db	373	GGGCGTGTCCCGGCGCTGCGCTGCTGTGCTGAACCGCGGGCGGCAAGGG	432
OY	258	CAAGGCTTTGAGCTCTTCGAGTACGTGACGCCCTTTGGCTGAGGCTGAATCTC	317
Db	433	CAAGGCTTTGAGCTCTTCGAGTACGTGACGCCCTTTGGCTGAGGCTGAATCTC	492
OY	318	CTTCAAGCTGATGCTCACTGAGCGGGAACACGCGGAGCTGTGCGTGGAGGA	377
Db	493	CTTCAAGCTGATGCTCACTGAGCGGGAACACGCGGAGCTGTGCGTGGAGGA	552
OY	378	GCTGGCGCGCTGGAGCGCTCT	398
Db	553	GCTGGCGCGCTGGAGCGCTCT	573

Search completed: July 12, 2003, 16:10:58
Job time : 382 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 15:18:22 ; Search time 4153 Seconds

(without alignments)
11205.246 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Sequence: 1599
1 ccccccggggctctctatagc.....aaataagtgacattcccaa 1599

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rnd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hugo_hum:*
40: em_hugo_mus:*
41: em_hugo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1600	6	AX224379 Sequence
2	623	39.0	1821	2	AK022402 Homo sapi
3	623	39.0	209861	2	AC021196 Homo sapi
4	590	36.9	1772	9	AK023393 Homo sapi
5	590	36.9	1783	9	AF238083 Homo sapi
6	590	36.9	1824	9	BC008040 Homo sapi
7	572	35.8	1192	9	BC004112 Homo sapi
8	572	35.8	1428	9	HSR245504 Homo sapi
9	572	35.8	1869	9	BC009419 Homo sapi
10	572	35.8	2502	9	AK095578 Homo sapi
11	524	32.8	1693	9	AF266756 Homo sapi
12	524	32.8	1719	6	AX127641 Homo sapi
13	521	32.6	183443	2	AC068145 Homo sapi
14	451	28.2	1173	9	AF200328 Homo sapi
15	433	27.1	1155	6	AX127642 Sequence
16	240	15.0	240	6	AB046025 Macaca fa
17	225	14.1	2015	9	AB046025 Sequence
18	192	12.0	296	6	AX198252 Sequence
19	192	12.0	296	6	AX208818 Sequence
20	191	11.9	394	6	AX334274 Sequence
21	167	10.4	62875	2	AC068436 Homo sapi
22	23	2.5	150350	9	CNS01DWV Human chr
23	25	1.6	1146	10	AF415213 Mus muscu
24	25	1.6	1559	6	AX287138 Sequence
25	25	1.6	1559	10	AF068749 Mus muscu
26	25	1.6	1759	6	AX224381 Sequence
27	25	1.6	1815	6	AX287137 Sequence
28	25	1.6	1815	10	AF068748 Mus muscu
29	25	1.6	110000	2	AL606505 Mus muscu
30	25	1.6	112037	10	AL645851 Mus muscu
31	23	1.4	109440	9	AC092891 Homo sapi
32	23	1.4	156339	2	AC068613 Homo sapi
33	22	1.4	70293	2	AC128914 Rattus no
34	22	1.4	102528	2	AL161779 Human DNA
35	22	1.4	106374	2	AC004390 Homo sapi
36	22	1.4	120937	2	AP004131 Oryza sat
37	22	1.4	123864	2	AP004859 Oryza sat
38	22	1.4	125348	2	AP004850 Oryza sat
39	22	1.4	177720	9	AC019193 Homo sapi
40	21	1.3	140	11	G54764 200H7L BAC
41	21	1.3	299	11	AF106665 Mus muscu
42	21	1.3	16713	1	AE006981 Mycobacte
43	21	1.3	24680	1	MTY21C12 Mycobacte
44	21	1.3	60933	9	AC112699 Homo sapi
45	21	1.3	152246	9	AC023473 Homo sapi

ALIGNMENTS

RESULT 1
AX224379
LOCUS AX224379 1600 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0160990.
ACCESSION AX224379
VERSION AX224379.1 GI:15554631
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Rastelli, L.
TITLE Novel sphingosine kinases
JOURNAL Patent: WO 0160990-A 1 23-AUG-2001;

Curagen Corporation (US) ; GENENTECH, INC. (US)
 Location/Qualifiers
 1.1600
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 265 a 492 c 531 g 311 t 1 others
 ORIGIN

Query Match 100.0%; Score 1599; DB 6; Length 1600;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

  1 CCCCCCGGGCTCTATAGCCAGGCGCTCCGGGGGGAAGGAGGCCACAGCCGCGCC 60
  2 CCCCCCGGGCTCTATAGCCAGGCGCTCCGGGGGGAAGGAGGCCACAGCCGCGCC 61
  61 TGGCAGCGCCGCTGGGCGAGCAGCATAGGAGCTGAAAGGCGAGAGCCCGCCAGCGGC 120
  62 TGGCAGCGCCGCTGGGCGAGCAGCATAGGAGCTGAAAGGCGAGAGCCCGCCAGCGGC 121
  121 AGCGCCCGCCAGCGCCAGGAGCCCGCTGGCAGGAGGAGCGCGGGTCAAGTTATGAT 180
  122 AGCGCCCGCCAGCGCCAGGAGCCCGCTGGCAGGAGGAGCGCGGGTCAAGTTATGAT 181
  181 CGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 240
  182 CGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 241
  241 CGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 300
  242 CGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 301
  301 GGTGAGGCTGAATCTCTTCAAGCTGATGCTCACTGAGCGGGGAGACCAAGCGGGAG 360
  302 GGTGAGGCTGAATCTCTTCAAGCTGATGCTCACTGAGCGGGGAGACCAAGCGGGAG 361
  361 CTGTGGGGTGGAGAGAGCTGGGGCGCTGGAGGCGTGGTGGTATCTTGGAGACGG 420
  362 CTGTGGGGTGGAGAGAGCTGGGGCGCTGGAGGCGTGGTGGTATCTTGGAGACGG 421
  421 CTGTGAGCAGAGGTGGTGAACGGGCTCATGAGAGCGGGCTGACTGGAGACCGCATCCAG 480
  422 CTGTGAGCAGAGGTGGTGAACGGGCTCATGAGAGCGGGCTGACTGGAGACCGCATCCAG 481
  481 AAGCCCTGTGTAGCTCCAGAGAGGCTGCGCAAGCGGCTGGACACTCTTGAACAT 540
  482 AAGCCCTGTGTAGCTCCAGAGAGGCTGCGCAAGCGGCTGGACACTCTTGAACAT 541
  541 TATGCTGGCTATGAGCAGGTACCAATGAAGACCTCTCTGACCACTGACGCTATGCTG 600
  542 TATGCTGGCTATGAGCAGGTACCAATGAAGACCTCTCTGACCACTGACGCTATGCTG 601
  601 TGGCGCCGGTGTCTCAACCATGAAACCTGTCTGTGACACGGCTTCGGGGCTGCGC 660
  602 TGGCGCCGGTGTCTCAACCATGAAACCTGTCTGTGACACGGCTTCGGGGCTGCGC 661
  661 TGGCTGTGTGTCTCAACCATGAAACCTGTCTGTGACACGGCTTCGGGGCTGCGC 720
  662 TGGCTGTGTGTCTCAACCATGAAACCTGTCTGTGACACGGCTTCGGGGCTGCGC 721
  721 AAGTATGGGCTGTGGGGAGATGCGCTTCACTGAGGACCTTCGGGCTGCGCAGCC 780
  722 AAGTATGGGCTGTGGGGAGATGCGCTTCACTGAGGACCTTCGGGCTGCGCAGCC 781
  781 CTGGCAGCTACCGGGCGGAGTGGCTACCTCCCTGTAGAGAGAGTGGGTTTCAAGACA 840
  782 CTGGCAGCTACCGGGCGGAGTGGCTACCTCCCTGTAGAGAGAGTGGGTTTCAAGACA 841
  841 CCGCTTCCCGCTTGTGTGTCAGAGAGGGCGCGGTATGACACACTGTGTGCGCACTGGAG 900
  842 CCGCTTCCCGCTTGTGTGTCAGAGAGGGCGCGGTATGACACACTGTGTGCGCACTGGAG 901
  901 GAGCAGGTGCTTCTCACTGAGGAGGTGTGCCGAGAGAGACTTGTCTAGTCTGAGCA 960
  
```

```

  902 GAGCAGGTGCTTCTCACTGAGGAGGTGTGCCGAGAGAGACTTGTGTAGTCTGAGCA 961
  961 CTGTGCTACCTGCGACCTGCGCAGTAGATGTTTGTCTGACCACTATGGCGCGCTGTGACGT 1020
  962 CTGTGCTACCTGCGACCTGCGCAGTAGATGTTTGTCTGACCACTATGGCGCGCTGTGACGT 1021
  1021 GGGCTCATGCTATGTTTCTACGTGGGGGGGAGTGTCTGCGCCATCTGCGGCTC 1080
  1022 GGGCTCATGCTATGTTTCTACGTGGGGGGGAGTGTCTGCGCCATCTGCGGCTC 1081
  1081 TTCTGTCAGTAGAGAGGCGCAGCATATGAGATGATGATGATGATGATGATGATGATG 1140
  1082 TTCTGTCAGTAGAGAGGCGCAGCATATGAGATGATGATGATGATGATGATGATGATG 1141
  1141 CCGGTGTGCTTCCGCTTGGAGCCCAAGAGATGAGAGAGTGTGTTGCACTGATGATG 1200
  1142 CCGGTGTGCTTCCGCTTGGAGCCCAAGAGATGAGAGAGTGTGTTGCACTGATGATG 1201
  1201 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
  1202 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1261
  1261 AGCGGTGCTGTGAGAGCCCGCCCGCAGAGTGAAGCCAGAGATGCGACCGCAGAGAG 1320
  1262 AGCGGTGCTGTGAGAGCCCGCCCGCAGAGTGAAGCCAGAGATGCGACCGCAGAGAG 1321
  1321 CCCTATGAGCCCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1380
  1322 CCCTATGAGCCCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1381
  1381 CCCTAGGGCTGAGAGGCGCTGTCCACAGTCTGTGGGGGGGGGAGAGAGTCTCTGAGAG 1440
  1382 CCCTAGGGCTGAGAGGCGCTGTCCACAGTCTGTGGGGGGGGGAGAGAGTCTCTGAGAG 1441
  1441 AGGGTGAAGAGTGAAGAGTATGCTTTGGGGGGGAGAGCCAGAAATGAAGTCTGGGCTAG 1500
  1442 AGGGTGAAGAGTGAAGAGTATGCTTTGGGGGGGAGAGCCAGAAATGAAGTCTGGGCTAG 1501
  1501 GAGCCCAAGCTGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1560
  1502 GAGCCCAAGCTGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1561
  1561 CAGCCCAAGCAATCCAAATGAATGAATGATGATGATGATGATGATGATGATGATG 1599
  1562 CAGCCCAAGCAATCCAAATGAATGAATGATGATGATGATGATGATGATGATGATG 1600
  
```

RESULT 2
 AK022402
 LOCUS
 DEFINITION
 Homo sapiens cDNA FLJ12340 f1s, clone MAMMA1002268, moderately similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
 AK022402
 AK022402.1 GI:10433790
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagaetsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Mutakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuno,Y., Nihomiya,K. and Iwayanagi,T.
 NEDO human cDNA sequencing project
 Unpublished
 JOURNAL
 REFERENCE
 2 (bases 1 to 1821)
 Isogai,T. and Otsuki,T.

TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEMO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'-6' 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES Location/Qualifiers
 source
 1. 1821
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MAMMA1002268"
 /tissue_type="Mammary gland"
 /clone_id="MAMMA1"
 /note="cloning vector: pME18SFL3"
 393..1547
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB14028.1"
 /db_xref="GI:10433791"
 /translation="MDPAGGPRGVLPAPRCRVLYLLNPRGKKAQLQFRSHVQPLAEAEISFTLMEFRNRHARELYSEELGRMDALVNSGDIEMHEVNGLMERDFTAIQKPLCSLPAGSGNALAASLNHYAGYEYNTNEDLLNCTLLCRLLSPMNLISHTAGLRLEFSYLSMGFTADYDSEKREKPEFTGCTRLAALTEGRLAYGVYGSRTPASPVYQGGPPYDAHLVPLEPVPSPHRTVYVDEDFYVLLALSHSGSEKPAAPMGRCAAGMHLFVIRAGVSRMLRLFLAMKGHMEECYLYVAVARLPKDGKGFVAVDEELVSAVOGVHPNFVNVSGVEPPSPKPOOMPPEPLP"
BASE COUNT 301 a 569 c 614 g 337 t
ORIGIN
 Query Match 39.0%; Score 623; DB 9; Length 1821;
 Best Local Similarity 99.9%; P-adj. No. 0;
 Matches 673; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

COMMENT
 Db 1465 TTGGGGGGAGCAGCCAGAAATGAAGTCCCTGGTCAGAGCCACAGCTGGCCAGCT 1524
 OY |||||||
 Db 1683 TTTGGGGGAGCAGCCAGAAATGAAGTCCCTGGTCAGAGCCACAGCTGGCCAGCT 1742
 OY |||||||
 Db 1525 GCCTATGTAGAGCCCTTCTAGTTTGTGTGAGACCCACCCACGAAACAAATCAATA 1584
 OY |||||||
 Db 1743 GCCTATGTAGAGCCCTTCTAGTTTGTGTGAGACCCACCCACGAAACAAATCAATA 1802
 OY |||||||
 Db 1585 AAGTACATTCCCA 1598
 OY |||||||
 Db 1803 AAGTACATTCCCA 1816
RESULT 3
 AC021196/c
 LOCUS
 DEFINITION Homo sapiens chromosome 17 clone RP11-794C22, WORKING DRAFT
 ACCESSION AC021196
 VERSION AC021196.3 GI:7631117
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
 TITLE The sequence of Homo sapiens clone
 AUTHORS Waterston, R.H.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 209861)
 AUTHORS Waterston, R.H.
 TITLE Submitted (14-JAN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 JOURNAL On Apr 21, 2000 this sequence version replaced g1:7577693.
 COMMENT
 ----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Project Information -----
 Center project name: H-NH0794C22
 ----- Summary Statistics -----
 Sequencing vector: M13; 91%
 Chemistry: Dye-Primer ET; 91% of reads
 Chemistry: Dye-terminator Big Dye; 9% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 188997 bases at least Q40
 Consensus quality: 195132 bases at least Q30
 Consensus quality: 198570 bases at least Q20
 Insert size: 225000; agarose-fp
 Insert size: 206861; sum-of-contigs
 Quality coverage: 3.50 in Q20 bases; agarose-fp
 Quality coverage: 3.75 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 31 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 2315: contig of 2315 bp in length
 * 2316 2415: gap of unknown length
 * 2416 4572: contig of 2157 bp in length
 * 4573 4673: gap of unknown length
 * 4673 7009: contig of 2337 bp in length
 * 7010 7109: gap of unknown length
 * 7109 9765: contig of 2656 bp in length
 * 9765 7110

Query Match	Best Local Similarity	Matches	673; Conservative	0; Mismatches	1; Indels	0; Gaps	0;
QY	925 GTGGTGC	CGCAGAGACTTGTGCTGACATGCTTCTTCTGCGCCATGAGAGGCGAGG	1104				
DB	72221 GTGGTGC	CGCAGAGACTTGTGCTGACATGCTTCTTCTGCGCCATGAGAGGCGAGG	1104				
QY	985 GAGATGTT	GGTGCACCATGAGGCGGTGTCAGACATGCTTCTTCTGCGCCATGAGAGGCGAGG	1104				
DB	72161 GAGATGTT	GGTGCACCATGAGGCGGTGTCAGACATGCTTCTTCTGCGCCATGAGAGGCGAGG	1104				
QY	1045 CGGCGGAG	TGCTGTCGATCCTCTGCGCCATGAGAGGCGAGG	1104				
DB	72101 CGGCGGAG	TGCTGTCGATCCTCTGCGCCATGAGAGGCGAGG	1104				
QY	1105 CATATGAG	TAAATGACCCCTACTTGCTATATGTCGCCGTGGTCCGCTTGGAG	1164				

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
CDS

Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1783)
Nava, V.E., Lacana, E., Poulton, S., Liu, H., Suglura, M., Kono, K.,
Milstien, S., Kohama, T. and Spiegel, S.
Functional characterization of human sphingosine kinase-1
FEBS Lett. 473 (1), 81-84 (2000)

2 (bases 1 to 1783)
Nava, V.E., Lacana, E., Poulton, S., Liu, H., Suglura, M., Kono, K.,
Milstien, S., Kohama, T. and Spiegel, S.
Direct Submission
Submitted (23-FEB-2000) Biochemistry, Georgetown University, 3900
Reservoir RD NW BSB Rm 357, Washington, DC 20007, USA

Location/Qualifiers
1..1783
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q25.2"
337..1491
/codon_start=1
/product="sphingosine kinase-1"
/protein_id="AAF73423.1"
/db_xref="GI:8132868"
/translation="MDPAGGRGVLPFRPCRYVLLNPRGKAKALQIFRSHVQPLAE
AEISPTLMTERNHARELYSEELRMDALVMSGDGLMHEVNGLMERPMETAIQ
KFLCSLPGSGNLAASLNHYAGYBOVNDLITNCTLLCRLLSPNMLSLHTASG
LRLFSVLSLMGFLADVDLSEKRYRLGEMRFTLGTLLALRLYRGRILAYLPVGRV
GSKTPASPVVVOQGVDAHLVPLEBPVSHMTVPDDEFLVLLALSHLSGSEFAAP
MGRCAAGVNHLEFVRAGVSRMLRLFLANERKRMHEKRYLYVYVVAFLRLEPKDG
KGFVAVDGEIWMSEAVOQGVHPTFMVWSGVCEPPPSKMPQOMPPEEPL"

BASE COUNT 314 a 546 c 591 g 332 t

Query Match 36.9%; Score 590; DB 9; Length 1783;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 590; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

18 AGCCAGGCTCCGGGCGGGAAGGAGCCGACAGCGGCGCTCGACGCCGCTGGG 77
|||||
180 AGCCAGGCTCCGGGCGGGAAGGAGCCGACAGCGGCGCTCGACGCCGCTGGG 239
|||||
78 CAGCAGGATAGAGAGCTGAAGGAGAGCCGCGGCGAGGCGGCGGCGGCGGCGG 137
|||||
240 CAGCAGGATAGAGAGCTGAAGGAGAGCCGCGGCGAGGCGGCGGCGGCGGCGG 299
|||||
138 AGGAGCCCGTGGAGGAGGAGCGGCGGCTGAGATTATGATCCAGCGGCGGCGGCGG 197
|||||
300 AGGAGCCCGTGGAGGAGGAGCGGCGGCTGAGATTATGATCCAGCGGCGGCGGCGG 359
|||||
198 GGGGCTGCTCCGCGCGCTGCGCGCTGCTGCTGCTGCAACCGCGGCGGCGGCGG 257
|||||
360 GGGGCTGCTCCGCGCGCTGCGCGCTGCTGCTGCTGCAACCGCGGCGGCGGCGG 419
|||||
258 CAGGCGCTTGGAGCTTTCGCGAGTCAAGTCAAGCCCTTTTGGCTGAGGCTGAATCTC 317
|||||
420 CAGGCGCTTGGAGCTTTCGCGAGTCAAGTCAAGCCCTTTTGGCTGAGGCTGAATCTC 479
|||||
318 CTTACAGCTGATGCTCACTAGAGCGGGAACCAAGCGGCGGAGCTGGTGGCGGAGGA 377
|||||
480 CTTACAGCTGATGCTCACTAGAGCGGGAACCAAGCGGCGGAGCTGGTGGCGGAGGA 539
|||||
378 GCTGGCGCGCTGGAGCGCTGCTGGTGCATGTCTGAGAGCGGCTGATGCACAGAGTGT 437
|||||
540 GCTGGCGCGCTGGAGCGCTGCTGGTGCATGTCTGAGAGCGGCTGATGCACAGAGTGT 599
|||||
438 GAACGGGCTCATGAGCGGCTGACTGGAGACCGCATCCAGAGCCCTGTGTACCT 497
|||||

Db 600 GAACGGGCTCATGAGCGGCTGACTGGAGAGCCGCGCATCCAGAGCCCTGTGTACCT 659
Qy 498 CCCAGAGGCTCTGGCAAGCGGCTGGAGCTTCTTGACCAATTATGCTGGCTATGACA 557
|||||
Db 660 CCCAGAGGCTCTGGCAAGCGGCTGGAGCTTCTTGACCAATTATGCTGGCTATGACA 719
|||||
Qy 558 GGTACCAATGAGAGCTCTGACCAACTGACAGCTATGCTGTGCGGCC 607
|||||
Db 720 GTTCACCAATGAGAGCTCTGACCAACTGACAGCTATGCTGTGCGGCC 769
|||||

RESULT 6
BC008040
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BC008040
Homo sapiens, clone MGC:15041 IMAGE:3831657, mRNA, complete cds.
BC008040.1 GI:14165485
MGC.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1824)
Strausberg, R.
Direct Submission
Submitted (21-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgaphs-remail.nih.gov
Tissue Procurement: ATCC/DOCT/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRML Plate: 24 Row: n Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10433790.
Location/Qualifiers
1..1824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:15041 IMAGE:3831657"
/tissue_type="Skin, melanotic melanoma."
/clone_1kb="NIH-MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOT7"
339..1493
/codon_start=1
/product="Unknown (protein for MGC:15041)"
/protein_id="AAH08040.1"
/db_xref="GI:14165486"
/translation="MDPAGGRGVLPFRPCRYVLLNPRGKAKALQIFRSHVQPLAE
AEISPTLMTERNHARELYSEELRMDALVMSGDGLMHEVNGLMERPMETAIQ
KFLCSLPGSGNLAASLNHYAGYBOVNDLITNCTLLCRLLSPNMLSLHTASG
LRLFSVLSLMGFLADVDLSEKRYRLGEMRFTLGTLLALRLYRGRILAYLPVGRV
GSKTPASPVVVOQGVDAHLVPLEBPVSHMTVPDDEFLVLLALSHLSGSEFAAP
MGRCAAGVNHLEFVRAGVSRMLRLFLANERKRMHEKRYLYVYVVAFLRLEPKDG
KGFVAVDGEIWMSEAVOQGVHPTFMVWSGVCEPPPSKMPQOMPPEEPL"

BASE COUNT 354 a 551 c 592 g 327 t

Query Match 36.9%; Score 590; DB 9; Length 1824;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 590; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 AGCCACGGCTCCGGGCGGGAGGCGAGCCACAGCCGCCCTGCGACGCCCTGGG 77
 Db 182 AGCCACGGCTCCGGGCGGGAGGCGAGCCACAGCCGCCCTGCGACGCCCTGGG 241
 QY 78 CAGCAGCATTAAGAGCTGAAGGAGAGCCGCCAGGAGGAGGAGCCGCCACAGGCC 137
 Db 242 CAGCAGCATTAAGAGCTGAAGGAGAGCCGCCAGGAGGAGGAGCCGCCACAGGCC 301
 QY 138 AGGAGACCCCTGGGAGAGGCGGGGCTGAGGTTATGATCCAGCGGGGCCCGG 197
 Db 302 AGGAGACCCCTGGGAGAGGCGGGGCTGAGGTTATGATCCAGCGGGGCCCGG 361
 QY 198 GGGGCTGCTCCCGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257
 Db 362 GGGGCTGCTCCCGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
 QY 258 CAGGACCTTGCAGCTCTTCCGAGTCAAGTCAAGCCCTTTTGGCTGAGGCTGAATCTC 317
 Db 422 CAGGACCTTGCAGCTCTTCCGAGTCAAGTCAAGCCCTTTTGGCTGAGGCTGAATCTC 481
 QY 318 CTTACGCTGATGCTCACTGAGCGCGGAGAACACAGCGCGGAGCTGGTGGGAGGA 377
 Db 482 CTTACGCTGATGCTCACTGAGCGCGGAGAACACAGCGCGGAGCTGGTGGGAGGA 541
 QY 378 GCTGGCGGCTGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437
 Db 542 GCTGGCGGCTGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
 QY 438 GAACGGCTCATGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
 Db 602 GAACGGCTCATGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
 QY 498 CCCAGCAGGCTTGGCAAGCGGCTGGAGCTTCTTGAACCATATGCTGGCTATGACA 557
 Db 662 CCCAGCAGGCTTGGCAAGCGGCTGGAGCTTCTTGAACCATATGCTGGCTATGACA 721
 QY 558 GGTACCAATGAAAGACTCTGACCAACTGACGACGCTATGCTGCGGCC 607
 Db 722 GGTACCAATGAAAGACTCTGACCAACTGACGACGCTATGCTGCGGCC 771

RESULT 7
 BC004112 1192 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, similar to sphingosine kinase 1, clone IMAGE:3832587,
 DEFINITION mRNA, partial cds.
 ACCESSION BC004112
 VERSION BC004112.1 GI:13278662
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1192)
 Strausberg, R.
 Direct Submission
 Submitted (01-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed By: The I.M.A.G.E. Consortium (ILNMF)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadnesystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNMF at: <http://image.llnl.gov>
 Series: IRAL Plate: 14 Row: f Column: 13.
 Location/Qualifiers

FEATURES
 SOURCE
 1. 1192
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3832587"
 /tissue_type="Skin, melanotic melanoma."
 /clone_id="NIH-MGC_20"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 <1. 873

CDS

/product="start=1
 /product="similar to sphingosine kinase 1"
 /protein_id="AA04112.1"
 /db_xref="GI:13278663"
 /translation="GTGTTAIOPLCSLPAAGSNAALASLNHYAGYEQVNTEDLTNC
 /translaltion="GTGTTAIOPLCSLPAAGSNAALASLNHYAGYEQVNTEDLTNC
 TLLCRLLSPMLLSLHSLAGSLFVSLANGFLVDLESEKYYRLGEMFTLGT
 FLRLAALRTYRGLAYLPVGRVGSKRTAGSVMYVVOGPDVADHLYLEPVSHTVVD
 EDPLVIALHSHLSGSEMFAPMGKASGVNHLFVYRAGVSRAMLRFLAMKGRHM
 EYECPRIVYTPVYAFRLPEPKDGKVFVDELNVSEAVQGVNPNFTFMVSGVEPFP
 SMKPQMPPEEP.
 BASE COUNT 245 a 345 c 357 g 245 t

ORIGIN

Query Match 35.8%; Score 572; DB 9; Length 1192;
 Best Local Similarity 99.7%; Pred. No. 1.3e-112;
 Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 925 GTGTGCCCCGAGAGAGCTTGTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 984
 Db 469 GTGTGCCCCGAGAGAGCTTGTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 528
 QY 985 GAGATGTTTGTGTCAGACCATGAGGCGGCTGTGCACTGCGCTGATGCACTGTTTACTG 1044
 Db 529 GAGATGTTTGTGTCAGACCATGAGGCGGCTGTGCACTGCGCTGATGCACTGTTTACTG 588
 QY 1045 CGGGCGGAGAGTGTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1104
 Db 589 CGGGCGGAGTGTGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
 QY 1105 CATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1164
 Db 649 CATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
 QY 1165 CCCAAGGATGGAAGAGTGTGTCAGTGTGATGAGGGAATGATGATGATGATGATGATG 1224
 Db 709 CCCAAGGATGGAAGAGTGTGTCAGTGTGATGAGGGAATGATGATGATGATGATGATG 768
 QY 1225 CAGGGCAGGTCAGCCCAACTACTTCTGATGATGATGATGATGATGATGATGATGAT 1284
 Db 769 CAGGGCAGGTCAGCCCAACTACTTCTGATGATGATGATGATGATGATGATGATGAT 828
 QY 1285 AGCTGGAAGCCCCAGAGATGCAACGCCAGAGAGAGCCCTTATGACCCCTGGGCGGCT 1344
 Db 829 AGCTGGAAGCCCCAGAGATGCAACGCCAGAGAGAGCCCTTATGACCCCTGGGCGGCT 888
 QY 1345 GTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1404
 Db 889 GTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 QY 1405 CAGCTCTGTTGGGGGTGAGAGAGCTCTGAGAGAGGAGGAGAGGAGGAGGAGGAGGAG 1464
 Db 949 CAGCTCTGTTGGGGGTGAGAGAGCTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1008
 QY 1465 TTTGGGGGAGACAGCCAGAAATGAAGTCTGTGGTCAAGAGCCAGCTGGTGGCCAGCT 1524
 Db 1009 TTTGGGGGAGACAGCCAGAAATGAAGTCTGTGGTCAAGAGCCAGCTGGTGGCCAGCT 1068
 QY 1525 GCTTATGATGAGGCTTCTAGTGTGTCAGAGAGCCAGCCAGCCAGCCAGCCAGCCAGCC 1584
 Db 1069 GCTTATGATGAGGCTTCTAGTGTGTCAGAGAGCCAGCCAGCCAGCCAGCCAGCCAGCC 1128

OY 1585 AAGTGCATTCCCA 1598
 |||||||
 Db 1129 AAGTGCATTCCCA 1142

 RESULT 8
 HSA245504 1428 bp mRNA linear PRI 22-MAY-2000
 LOCUS Homo sapiens mRNA for sphingosine kinase (gene SPK).
 DEFINITION A245504
 ACCESSION A245504.1 GI:8017375
 VERSION sphingosine kinase; SPK gene.
 KEYWORDS sphingosine kinase; SPK gene.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1428)
 AUTHORS Van Veldhoven, P.P. and Gijbers, S.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1428)
 AUTHORS Van Veldhoven, P.P.
 TITLE Direct Submission
 JOURNAL Submitted (16-AUG-1999) Van Veldhoven, P.P., Molecular Cell Biology,
 Katholieke Universiteit Leuven, K.U. Leuven, Campus Gasthuisberg,
 Afid. Farmakologie, Herestraat, B-3000 Leuven, Belgium
 REVISION Revised by author (17-MAY-2000)
 FEATURES
 source location/Qualifiers
 1..1428
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q24"
 1..1146
 /gene="SPK"
 1..1146
 /gene="SPK"
 /codon_start=1
 /product="sphingosine kinase"
 /protein_id="CAB92131.1"
 /db_xref="GI:8017376"
 /translation="MDPAGRGVARLRVLLINPRGKGLQLFRSHVPLAEEL
 STPLMTERNRHARLSEELREMDALVMSGDLMHEVNGMLERDFTALQKPL
 CTPASGSKPLASLNHAYGBOVTNEDLLCTLLCRLLSPNNLLSHTAGLRL
 FSYLSANFPIADVDLESERKRLGEMRPTLGTFLALATYRGLAYLPGVGRSK
 TPASPVVQDGPVDALVPLERPVSHMTVVDDEEVYLLHLSGSEMFAPMGR
 CAAGVHLFVRAGVSRNALRLFLAMERGRMECEPPIVYVPGVGRLEPKDKGV
 FAVDGLWSEAVQGVHPNFMVSGVCEPPEPPSMKPPQMPPEEPL"
 BASE COUNT 247 a 422 c 462 g 297 t
 ORIGIN
 Query Match 35.8%; Score 572; DB 9; Length 1428;
 Best Local Similarity 99.7%; Pred. No. 1,3e-312;
 Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 982 CCCAAGATGGAAAGGTGTGTTTCAGTGGATGGGGAATGATGTTAGCGAGGCCGTG 1041
 OY 1225 CAGGGCCAGGTGCACCCAACTACTTGTGATGGTGCAGGGGTGCTGAGAGCCCGCC 1284
 |||||||
 Db 1042 CAGGGCCAGGTGCACCCAACTACTTGTGATGGTGCAGGGGTGCTGAGAGCCCGCC 1101
 OY 1285 AGCTGGAGAGCCCGACAGATGCACCGCCAGAAAGCCCTTATGACCCCTGGCGGCT 1344
 |||||||
 Db 1102 AGCTGGAGAGCCCGACAGATGCACCGCCAGAAAGCCCTTATGACCCCTGGCGGCT 1161
 OY 1345 GTGCTTAACTGTCTACTTGCAGAGCCCTTCCTCTTCTAGGGGTGAGGGCCGTGCA 1404
 |||||||
 Db 1162 GTGCTTAACTGTCTACTTGCAGAGCCCTTCCTCTTCTAGGGGTGAGGGCCGTGCA 1221
 OY 1405 CAGCTCTGTGGGGGTGAGAGACTCTCTGTGAGAGAGGTGAGAGGTGAGGTATGC 1464
 |||||||
 Db 1222 CAGCTCTGTGGGGGTGAGAGACTCTCTGTGAGAGAGGTGAGAGGTGAGGTATGC 1281
 OY 1465 TTTGGGGGAGCAGGCCAGAAATGAATGATCTCTGAGAGAGCCAGGCTGGGCGCAGCT 1524
 |||||||
 Db 1282 TTTGGGGGAGCAGGCCAGAAATGAATGATCTCTGAGAGAGCCAGGCTGGGCGCAGCT 1341
 OY 1525 GCGTATGTAAGGCGCTTCTAGTTGTTCTGAGAGCCCGCCAGCCAGCAACCAATCCAAATA 1584
 |||||||
 Db 1342 GCGTATGTAAGGCGCTTCTAGTTGTTCTGAGAGCCCGCCAGCCAGCAACCAATCCAAATA 1401
 OY 1585 AAGTGACATTCCCA 1598
 |||||||
 Db 1402 AAGTGACATTCCCA 1415

 RESULT 9
 BC009419 1869 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, clone MGC:15040 IMAGE:3506924, mRNA, complete cds.
 DEFINITION BC009419
 ACCESSION BC009419
 VERSION BC009419.1 GI:14495624
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1869)
 AUTHORS Mammalian
 TITLE Direct Submission
 JOURNAL Submitted (19-JUN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chiu, Readman Chiu, Chris Fjell, Erin Garland, Kan Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Maitra.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRL Plate: 23 Row: e Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 10433790.
 Location/Qualifiers

[illegible]

Db	1833	AAGTACATTCCTCCA	1846
RESULT_10			
LOCUS	AK095578	2502 bp	mRNA linear PRI 15-JUL-2002
DEFINITION	Homo sapiens cDNA FLJ38259, f1s, clone FCBBF3001302, highly similar to Homo sapiens sphingosine kinase (SPHK1) mRNA.		
ACCESSION	AK095578		
VERSION	AK095578.1	GI:21754862	
SOURCE	Oligo capping; f1s (full insert sequence).		
ORGANISM	Homo sapiens fetal brain CDNA to mRNA, clone_11b:FCBBF3 clone:FCBBF3001302.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Salto, K., Nishikawa, T., Kimura, K., Yanashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Magatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K., Masuhio, Y., Negai, K. and Isogal, T.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2502)		
AUTHORS	Isogal, T. and Yamamoto, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-JUL-2002) Takao Isogal, FUJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan		
COMMENT	(E-mail:genomics@hri.co.jp, Tel:01-438-52-3975, Fax:01-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.		
FEATURES	Location/Qualifiers		
SOURCE	1..2502		
	/organism="Homo sapiens"		
	/db_xref="taxon:3606"		
	/clone="FCBBF3001302"		
	/tissue_type="brain"		
	/clone_11b="FCBBF3"		
	/dev_stage="fetal"		
	/note="cloning vector: PME18SF13"		
BASE COUNT	404 a 803 c 835 g 460 t		
ORIGIN			
Query Match	35.8%; Score 572; DB 9; Length 2502;		
Best Local Similarity	99.7%; Pred. No. 1.3e-312;		
Matches 672:	Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY	925 GTGGTGCCCGCAGGACTTTGTGTGAATCCTGGCACTGCTGCACACTGCCACCTGGCCAGT	984	
Db	1828 GTGGTGCCCGCAGGACTTTGTGTGAATCCTGGCACTGCTGCACACTGCCACCTGGCCAGT	1887	
QY	985 GAGATGTTGGTGGCACCCATGGGCGCGTGTGCAGCTGGCGCATGCATCTGTTCACGTG	1044	
Db	1888 GAGATGTTGGTGGCACCCATGGGCGCGTGTGCAGCTGGCGCATGCATCTGTTCACGTG	1947	
QY	1045 CGGGCGGAGTGTCTGTGCATGCTGCGGCTCTTCGAGCCATGAGAGAAGGCGAGG	1104	
Db	1948 CGGGCGGAGTGTCTGTGCATGCTGCGGCTCTTCGAGCCATGAGAGAAGGCGAGG	2007	
QY	1105 CATANGAGTAGATGCCCCCTACTGGTATTATAGTGGCCGCGTGCCTTCGCGTTGGAG	1164	
Db	2008 CATANGAGTAGATGCCCCCTACTGGTATTATAGTGGCCGCGTGCCTTCGCGTTGGAG	2067	

FEATURES	source	Location/Qualifiers
BASE COUNT	305 a 529 c 556 g 329 t	
ORIGIN		
Query Match	Best Local Similarity 99.78% Pident. No. 2.4e-285; Matches 624; Conservative 0; Mismatch 2; Indels 0; Gaps 0;	
OY	925 GTGGTGCCTCCGACGAGCACTTTGTGTCTATGTCTTGCGCAGCTGTCAGCTGCACACTGCCACTGGCCAGT	984
Db	1020 GTGGTGCCTCCGACGAGCACTTTGTGTCTATGTCTTGCGCAGCTGTCAGCTGCACACTGCCACTGGCCAGT	1079
OY	985 GAGATGTTTTGCTGCACCCTCATGGGGCCGCTGTGTGACCTGGCCGTATGCATCTGTTCTACGTG	1044
Db	1080 GAGATGTTTTGCTGCACCCTCATGGGGCCGCTGTGTGACCTGGCCGTATGCATCTGTTCTACGTG	1139
OY	1045 CGGGCGGGAGTGTCTCTGTGCCATGTCTGTGGCCCTCTTCTCGGCCATGGAAGAAGGACAG	1104
Db	1140 CGGGCGGGAGTGTCTCTGTGCCATGTCTGTGGCCCTCTTCTCGGCCATGGAAGAAGGACAG	1199
OY	1105 CATATGAGATATGATGGCCCTACTGTGGTATATGGCCCCGGTGGCCCTTCCGTTGGAG	1164
Db	1200 CATATGAGATATGATGGCCCTACTGTGGTATATGGCCCCGGTGGCCCTTCCGTTGGAG	1259
OY	1165 CCCAAGGATGGGAAAGGTGTGTTTGACAGTGGATGGGAAATTGATGTTAGCGAGCCGTG	1224
Db	1260 CCCAAGGATGGGAAAGGTGTGTTTGACAGTGGATGGGAAATTGATGTTAGCGAGCCGTG	1319
OY	1225 CAGGGCCAGGTGTGACCCCAAATACTTCTGGATGTGTGACAGCGGTGTGCTGTGAGACCCCCTGCC	1284
Db	1320 CAGGGCCAGGTGTGACCCCAAATACTTCTGGATGTGTGACAGCGGTGTGCTGTGAGACCCCCTGCC	1379
OY	1285 AGCTGGAAGCCCCACAGCATGCCACCGCACAGAACGCCCTTATGACCCCTGTGGCGGAGCT	1344
Db	1380 AGCTGGAAGCCCCACAGCATGCCACCGCACAGAACGCCCTTATGACCCCTGTGGCGGAGCT	1439
OY	1345 GTGCTTAGTGTCTACTTGTGACGAGACCTTCTCTCTTCCCTTAAGGGCTGTGAGGGCTGTACA	1404
Db	1440 GTGCTTAGTGTCTACTTGTGACGAGACCTTCTCTCTTCCCTTAAGGGCTGTGAGGGCTGTACA	1499
OY	1405 CAGCTCCTGTGTGGGGGTGAGAGACCTCTCTGTGGAAGAGGTGAGAGGTGAGAGCTATGC	1464
Db	1500 CAGCTCCTGTGTGGGGGTGAGAGACCTCTCTGTGGAAGAGGTGAGAGGTGAGAGCTATGC	1559
OY	1465 TTGAGGGGGACAGGACGAGATGAAGTCTCTGGGTGAGGAGCCACAGCTGGCTGGGCCAGCT	1524
Db	1560 TTGAGGGGGACAGGACGAGATGAAGTCTCTGGGTGAGGAGCCACAGCTGGCTGGGCCAGCT	1619
OY	1525 GCCTATGTGAAGGCTTCTAGTTTGT	1550
Db	1620 GCCTATGTGAAGGCTTCTAGTTTGT	1645
RESULT 13		
AC068145/c	183443 bp DNA linear HTG 04-MAY-2001	
LOCUS	Homo sapiens chromosome 17 clone CTD-2531H7 map 17, WORKING DRAFT	
DEFINITION	SEQUENCE, 48 unordered pieces.	
ACCESSION	AC068145	
VERSION	AC068145.3 GI:9966946	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	Barren,B., Linton,L., Nusbaum,C. and Landier,E.	
TITLE	Homo sapiens chromosome 17, clone CTD-2531H7	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 183443)	

AUTHOR'S
 Arrhen, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F.,
 Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
 Campoliano, A., Castle, A., Chello, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., DeAngelio, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
 Gajana, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hages, B., Heatford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Kardas, A.,
 Klein, J., LaRoque, K., Lamatas, R., Landers, T., Lehotzky, J.,
 Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McNetters, R.,
 Melidrim, J., Menues, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisanli, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J.
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted 129-APR-2000 Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 5, 2000 this sequence version replaced gl:1960325.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: IL0026
 Center clone name: 2531.H.7
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 99% of reads
 Sequencing vector: Plasmid; n/a; %0.1% of reads
 0.756725341426404Chemistry: Dye-terminator Big Dye; 100% of
 reads

 Assembly program: Phrap; version 0.960731
 Consensus quality: 152281 bases at least Q40
 Consensus quality: 167680 bases at least Q30
 Consensus quality: 174098 bases at least Q20
 Insert size: 200000; agarose-fp
 Insert size: 178743; sum-of-confits
 Quality coverage: 2.9 in Q20 bases; agarose-fp
 Quality.
 NOTE: This is a 'working draft' sequence. It currently
 consists of 48 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 * 1 1059: contig of 1059 bp in length
 * 1060 1159: gap of 100 bp
 * 1160 2777: contig of 1618 bp in length
 * 2778 2877: gap of 100 bp
 * 2878 4283: contig of 1406 bp in length
 * 4284 4383: gap of 100 bp
 * 4384 5447: contig of 1064 bp in length
 * 5448 5547: gap of 100 bp
 * 5548 6776: contig of 1223 bp in length
 * 6777 6876: gap of 100 bp
 * 6877 8215: contig of 1339 bp in length
 * 8216 8315: gap of 100 bp
 * 8316 9961: contig of 1646 bp in length
 * 9962 10061: gap of 100 bp
 * 10062 12224: contig of 2163 bp in length
 * 12225 12324: gap of 100 bp

13325	13794:	contlg of 1470 bp	in length
13795	13894:	gap of 100 bp	
13895	16887:	contlg of 2963 bp	in length
16888	16957:	gap of 100 bp	
16958	18266:	contlg of 1309 bp	in length
18267	18366:	gap of 100 bp	
18367	20130:	contlg of 1824 bp	in length
20191	20290:	gap of 100 bp	
20291	21460:	contlg of 1170 bp	in length
21461	21560:	gap of 100 bp	
21561	23407:	contlg of 1847 bp	in length
23408	23507:	gap of 100 bp	
23508	25455:	contlg of 1948 bp	in length
25456	25555:	gap of 100 bp	
25556	27331:	contlg of 1776 bp	in length
27332	27441:	gap of 100 bp	
27432	30406:	contlg of 2975 bp	in length
30407	30506:	gap of 100 bp	
30507	32805:	contlg of 2299 bp	in length
32806	32905:	gap of 100 bp	
32906	33655:	contlg of 2754 bp	in length
33660	35755:	gap of 100 bp	
35760	37974:	contlg of 2215 bp	in length
37975	38074:	gap of 100 bp	
38075	40521:	contlg of 2447 bp	in length
40522	40621:	gap of 100 bp	
40622	43219:	contlg of 2598 bp	in length
43220	43319:	gap of 100 bp	
43320	44161:	contlg of 2842 bp	in length
44162	46261:	gap of 100 bp	
46262	48463:	contlg of 2202 bp	in length
48464	48563:	gap of 100 bp	
48564	51275:	contlg of 2712 bp	in length
51276	51375:	gap of 100 bp	
51376	53522:	contlg of 2147 bp	in length
53523	53622:	gap of 100 bp	
53623	56955:	contlg of 3333 bp	in length
56956	57055:	gap of 100 bp	
57056	59555:	contlg of 2500 bp	in length
59556	59655:	gap of 100 bp	
59656	62713:	contlg of 3058 bp	in length
62714	62813:	gap of 100 bp	
62814	66519:	contlg of 3706 bp	in length
66520	66619:	gap of 100 bp	
66620	68930:	contlg of 2311 bp	in length
68931	69030:	gap of 100 bp	
69031	72724:	contlg of 3694 bp	in length
72725	72824:	gap of 100 bp	
72825	76092:	contlg of 3268 bp	in length
76093	76192:	gap of 100 bp	
76193	80143:	contlg of 3951 bp	in length
80144	80243:	gap of 100 bp	
80244	84957:	contlg of 4714 bp	in length
84958	85057:	gap of 100 bp	
85058	88444:	contlg of 3367 bp	in length
88452	88524:	gap of 100 bp	
88525	93448:	contlg of 4904 bp	in length
93449	93538:	gap of 100 bp	
93539	97117:	contlg of 3649 bp	in length
97178	97217:	gap of 100 bp	
97278	102713:	contlg of 5436 bp	in length
102714	102813:	gap of 100 bp	
102814	109141:	contlg of 6328 bp	in length
109142	109241:	gap of 100 bp	
109242	114513:	contlg of 5272 bp	in length
114514	114613:	gap of 100 bp	
114614	124525:	contlg of 9912 bp	in length
124526	124625:	gap of 100 bp	
124626	132364:	contlg of 7739 bp	in length
132365	132464:	gap of 100 bp	
132465	133486:	contlg of 7024 bp	in length
133487	139588:	gap of 100 bp	
139589	148734:	contlg of 9146 bp	in length

FEATURES	SOURCE
*	148735 148834: gap of 100 bp
*	148835 159465: contig of 10631 bp in length
*	159466 159565: gap of 100 bp
*	159566 172685: contig of 13120 bp in length
*	172686 172785: gap of 100 bp
*	172786 183443: contig of 10658 bp in length
	Location/Qualifiers
	1..183443

Query Match	32.6%	Score 521;	DB 2;	Length 183443;
Best Local Similarity	99.68%	Pred. No. 1.3e+283;		
Matches 671;	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;

```
Db 13420 GTGCTTAGTGTACTTGGACGACCTTCCTCTCCCTAGAGCTGCAGAGGCTGTCCCA 133361
Qy 1405 CAGTCTCTGTGGGGGTGGAGAGACTCTCTGGAGAAAGGTGAGAGCTGAGCTATGC 1464
Db 13360 CAGTCTCTGTGGGGGTGGAGAGACTCTCTGGAGAAAGGTGAGAGCTGAGCTATGC 133301
Qy 1465 TTGGGGGGGACAGGCGCAAGTGAAGTCTGGTGGAGAGCCAGCTGGTGGGCCAGCT 1524
Db 13300 TTGGGGGGGACAGGCGCAAGTGAAGTCTGGTGGAGAGCCAGCTGGTGGGCCAGCT 133241
Qy 1525 GCCATATGAAGCCCTTCTAGTTTGTCTGAGACCCCAACCCAGCAACAATCAATA 1584
Db 133240 GCCATATGAAGCCCTTCTAGTTTGTCTGAGACCCCAACCCAGCAACAATCAATA 133181
Qy 1585 AAGTGACATTCCCA 1598
Db 133180 AAGTGACATTCCCA 133167

RESULT 14
AF200328 1173 bp mRNA linear PRI 25-AUG-2000
LOCUS Homo sapiens sphingosine kinase (SPHK) mRNA, complete cds.
DEFINITION AF200328
ACCESSION AF200328
VERSION AF200328.1 GI:9909360
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1173)
AUTHORS Pitson,S.M., D'andrea,R.J., Vandeleur,L., Moretti,P.A., Xia,P.,
Gamble,J.R., Vadas,M.A. and Wattenberg,B.W.
TITLE Human sphingosine kinase: purification, molecular cloning and
characterization of the native and recombinant enzymes
JOURNAL Biochem. J. 350 Pt 2, 429-441 (2000)
MEDLINE 20407120
PUBMED 10947957
2 (bases 1 to 1173)
Pitson,S.M., D'andrea,R.J., Vandeleur,L., Moretti,P.A.B., Xia,P.,
Gamble,J.R., Vadas,M.A. and Wattenberg,B.W.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-1999) Human Immunology, IMVS, Frome Road,
Adelaide, SA 5000, Australia
FEATURES
Source Location/Qualifiers
1..1173
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="vascular endothelial"
1..1173
/gene="SPHK"
19..1173
/gene="SPHK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="AA601980.1"
/db_xref="GI:9909361"
/translating="MDPAGPRGVLPFRCAVLLINRGGKQALQFRSHVQBLAE
AEISFTLMEERHARELVRSEBELGMDLVVMSGGGLNHEVNGMLERDMSHTAIO
KPLCSLPAGSGLNLAASLNHAGYEQVTEDELLNCLLCRLSPMNLISHTAG
LRLEFVSLAMGFIADVDLSEKRYRLEKRFITGFLRLAALRTYGRRLAYLPVGAV
GSKRPASPVVVOGPDVAHLVPLEEPYPSMTYVPPDEPVLVLAALSHSGSEKFAAP
MGRCAAVVHLFTYRACVSRAMLRILFLAEKGRHMETECRYLYTVVAVFRLEPKDG
KGMFADVELSVSAVGOVHPNTEFNVVSCVEPPSMKPDQMPPEPL"
BASE COUNT 188 a 349 c 390 g 246 t
ORIGIN
Query Match 28.2% Score 451; DB 9; Length 1173;
Best Local Similarity 100.0%; Pred. No. 6.9e-244;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 1 GAGCGGGGGTCAGAGTTATGATCCAGCGGGGGGCCCGGGGGCGTGTCCCGGGCCC 60
Qy 217 TGCCGCTGCTGCTGCTGCTGAACCCCGCGGGCGGCAAGGCAAGGCTTTC 276
Db 61 TGCCGCTGCTGCTGCTGCTGAACCCCGCGGGCGGCAAGGCAAGGCTTTC 120
Qy 277 CGGAGTACAGGCGGCCCTTTGGCTGAGAGCTGAATTCCTTACAGCTGATGCTACT 336
Db 121 CGGAGTACAGGCGGCCCTTTGGCTGAGAGCTGAATTCCTTACAGCTGATGCTACT 180
Qy 337 GAGCGGGGAAACACAGCGGGGAGAGCTGTGCGGTGGAGAGAGCTGGGCCCTGGAGAGCT 396
Db 181 GAGCGGGGAAACACAGCGGGGAGAGCTGTGCGGTGGAGAGAGCTGGGCCCTGGAGAGCT 240
Qy 397 CTGGTGTGATGCTGTGAGAGAGCGGCTGATGACAGAGGTGTGAACGGGCTCATAGACGG 456
Db 241 CTGGTGTGATGCTGTGAGAGAGCGGCTGATGACAGAGGTGTGAACGGGCTCATAGACGG 300
Qy 457 CCTGACTGGGAGAGACCGGCATCCAGAACCCCTGTGTAGCTTCCAGAGGCTCTGGGAC 516
Db 301 CCTGACTGGGAGAGACCGGCATCCAGAACCCCTGTGTAGCTTCCAGAGGCTCTGGGAC 360
Qy 517 GCGCTGCAGCTTCTTGAACCATATTATGTGCTATGAGAGAGCTCACCAATGAGAGCTC 576
Db 361 GCGCTGCAGCTTCTTGAACCATATTATGTGCTATGAGAGAGCTCACCAATGAGAGCTC 420
Qy 577 CTGACCAACTGACGCTATTGCTGTGCCGCC 607
Db 421 CTGACCAACTGACGCTATTGCTGTGCCGCC 451

RESULT 15
AX127642 1155 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 2 from Patent WO0131029.
DEFINITION AX127642
ACCESSION AX127642
VERSION AX127642.1 GI:14134308
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1155)
AUTHORS Allen,J., Gosink,M., Melendez,A.J. and Takacs,L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 2 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
Source Location/Qualifiers
1..1155
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 185 a 349 c 382 g 239 t
ORIGIN
Query Match 27.1% Score 433; DB 6; Length 1155;
Best Local Similarity 100.0%; Pred. No. 1.1e-233;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Db      181  CGGGAGCTGGTCGTCGAGAGAGCTGGCCGCTGGGACGCTCTGGTGTGATGTCTGGA 240
OY      415  GACGGGCTGATGTACAGAGTGTGAACGGGCTTCATGAGGCGCTGACTGGGAGACCGCC 474
Db      241  GACGGGCTGATGTACAGAGTGTGAACGGGCTTCATGAGGCGCTGACTGGGAGACCGCC 300
OY      475  ATCCAGAGCCCTGTGTAGCTCCAGCAGGCTCTGGCAACGCGCTGGCAGCTTCCTTG 534
Db      301  ATCCAGAGCCCTGTGTAGCTCCAGCAGGCTCTGGCAACGCGCTGGCAGCTTCCTTG 360
OY      535  AACCATATGCTGCTATGAGCAGTCCACCAATGAGAGCCTCTGACCAACTGCACGCTA 594
Db      361  AACCATATGCTGCTATGAGCAGTCCACCAATGAGAGCCTCTGACCAACTGCACGCTA 420
OY      595  TTGCTGTGCCGCC 607
Db      421  TTGCTGTGCCGCC 433

```

Search completed: July 12, 2003, 17:20:21
 Job time : 4156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 14:02:42 ; Search time 273 Seconds

(Without alignments)
9248,401 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599
Sequence: 1 cccccggggctctctatagc.....aaataagtgacattcccaa 1599

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1105431 seqs, 789497651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PC7_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PC7US_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1599	100.0	1600	US-09-784-810A-1	Sequence 1, Appl1
2	1426.8	89.2	1533	US-10-023-282-90	Sequence 90, Appl1
3	1124.6	70.3	1155	US-09-970-516-1	Sequence 1, Appl1
4	836.4	52.3	1759	US-09-784-810A-3	Sequence 3, Appl1
5	742.4	46.4	1149	US-09-970-516-5	Sequence 5, Appl1
6	382.6	23.9	394	US-09-954-456-1756	Sequence 1756, Ap
7	286.6	17.9	480	US-09-783-590-9248	Sequence 9248, Ap
8	258	16.1	1857	US-09-970-516-3	Sequence 3, Appl1
9	258	16.1	2380	US-09-817-676A-13	Sequence 13, Appl1
10	242	15.1	2698	US-09-817-676A-11	Sequence 11, Appl1
11	197.4	12.3	199	US-09-796-692-2905	Sequence 2905, Ap
12	197.4	12.3	199	US-10-040-862-2905	Sequence 2905, Ap
13	193	12.1	296	US-10-015-219-658	Sequence 658, App
14	193	12.1	296	US-09-777-564-658	Sequence 658, App
15	51.8	3.2	2136	US-10-156-761-7476	Sequence 7476, Ap
16	51.8	3.2	3135	US-10-156-761-5936	Sequence 5936, Ap
17	51.8	3.2	9025608	US-10-156-761-1	Sequence 1, Appl1
18	51.8	3.2	9025608	US-10-156-761-1	Sequence 1, Appl1
19	51.2	3.2	1614	US-09-969-896-9	Sequence 9, Appl1

20	51.2	3.2	4413	US-09-969-896-16	Sequence 16, Appl1
21	49.6	3.1	1203	US-09-879-312-1	Sequence 1, Appl1
22	47.6	3.0	861	US-10-156-761-2134	Sequence 2134, Ap
23	47.6	2.9	68750	US-10-014-717-1	Sequence 1, Appl1
24	46.6	2.9	792	US-10-156-761-3938	Sequence 3938, Ap
25	46.6	2.9	1734	US-10-156-761-4994	Sequence 4994, Ap
26	46.4	2.9	2949	US-10-187-267A-12	Sequence 12, Appl1
27	46.4	2.9	36321	US-10-187-267A-1	Sequence 1, Appl1
28	46.4	2.9	53522	US-09-904-968A-1	Sequence 1, Appl1
29	46.2	2.9	11817	US-10-156-761-2884	Sequence 2884, Ap
30	46.2	2.9	125746	US-10-156-761-15102	Sequence 15102, A
31	46	2.9	1026	US-10-156-761-3671	Sequence 3671, Ap
32	46	2.9	3834	US-10-156-761-2681	Sequence 2681, Ap
33	45.8	2.9	1542	US-10-156-761-7255	Sequence 7255, Ap
34	45.6	2.9	1155	US-10-156-761-4992	Sequence 4992, Ap
35	45.6	2.9	4935	US-10-156-761-5059	Sequence 5059, Ap
36	45.6	2.9	9968	US-09-764-877-2718	Sequence 2718, Ap
37	45.4	2.8	1092	US-10-156-761-1710	Sequence 1710, Ap
38	45.4	2.8	2808	US-10-156-761-2021	Sequence 2021, Ap
39	45.2	2.8	3135	US-10-156-761-5936	Sequence 5936, Ap
40	45.2	2.8	158405	US-10-175-523-86	Sequence 86, Appl
41	45	2.8	1389	US-10-156-761-2751	Sequence 2751, Ap
42	45	2.8	1392	US-10-156-761-2601	Sequence 2601, Ap
43	44.8	2.8	1143	US-10-126-279-22	Sequence 22, Appl
44	44.8	2.8	1542	US-10-156-761-4659	Sequence 4659, Ap
45	44.8	2.8	1614	US-09-976-740-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-09-784-810A-1
Sequence 1, Application US/09784810A
Patent No. US20020082203A1
GENERAL INFORMATION:
APPLICANT: RASTELLI, LUCA
TITLE OF INVENTION: NOVEL SPHINGOSINE KINASES AND NUCLEIC ACIDS ENCODING
FILE REFERENCE: 10/16-08
CURRENT APPLICATION NUMBER: US/09/784, 810A
CURRENT FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 1600
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (1)
OTHER INFORMATION: a, t, c, g, other or unknown
US-09-784-810A-1
Query Match 100.0%; Score 1599; DB 10; Length 1600;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCCCCGGGGCTCTATAGCCAGGCTCCGGGGGGAAGCCGACACCCACAGCCGGCC 60
2 CCCCCGGGGCTCTATAGCCAGGCTCCGGGGGGAAGCCGACACCCACAGCCGGCC 61
61 TGGCAGCGCCCTCTGAGCAGCAGCATTAAGAGCTGAAGAGCAGAGCCGCCAGCGGC 120
62 TGGCAGCGCCCTCTGAGCAGCAGCATTAAGAGCTGAAGAGCAGAGCCGCCAGCGGC 121
121 AGCGCCCGCCAGAGCGCCAGGACCCCTGAGCAGCGGAGCGCGGCTGAGGTATGAT 180
122 AGCGCCCGCCAGAGCGCCAGGACCCCTGAGCAGCGGAGCGCGGCTGAGGTATGAT 181

QY	181	CCAGCGGGGCGGCCCCCGGAGCGTGTCTCCCGGGCCCTGCGCGGTGGGTGGCTGAAC	240
Db	182	CCAGCGGGGCGGCCCCCGGAGCGTGTCTCCCGGGCCCTGCGCGGTGGGTGGCTGAAC	241
QY	241	CCGCGGGGGGCGAAGGCGAAGGCTTTGACGCTTTTCCGGATCAGCTGACGCCCTTTTG	300
Db	242	CCGCGGGGGGCGAAGGCGAAGGCTTTGACGCTTTTCCGGATCAGCTGACGCCCTTTTG	301
QY	301	GCTGAGGCTGAATCTCTTTCACGCTGATGCTCACTAGAGGGCGGCGAACCAGCGCGAG	360
Db	302	GCTGAGGCTGAATCTCTTTCACGCTGATGCTCACTAGAGGGCGGCGAACCAGCGCGAG	361
QY	361	CTGTGTGGGCTCGGAGAGAGCTGGGCGCGTGGAGCGCTCTGTGTGTATGTCTGGAGAGCGG	420
Db	362	CTGTGTGGGCTCGGAGAGAGCTGGGCGCGTGGAGCGCTCTGTGTGTATGTCTGGAGAGCGG	421
QY	421	CTGATGACAGAGGTGGGGAACGGGCTCATGAGAGCGGCTGACTGTGGAGACCGCCATCAG	480
Db	422	CTGATGACAGAGGTGGGGAACGGGCTCATGAGAGCGGCTGACTGTGGAGACCGCCATCAG	481
QY	481	AAGCCCGCTGTAGCCTCCAGAGGCTCTGGACACGCGCTGGGAGGCTTCTTGAACCAT	540
Db	482	AAGCCCGCTGTAGCCTCCAGAGGCTCTGGACACGCGCTGGGAGGCTTCTTGAACCAT	541
QY	541	TATGTGTGCTATGAGCAGCTACCAATGAGACCTCTGACCACTGCACGCTATTGCTG	600
Db	542	TATGTGTGCTATGAGCAGCTACCAATGAGACCTCTGACCACTGCACGCTATTGCTG	601
QY	601	TGCGCGCCGCTGTCTACACCATGAAACCTCTCTCTGTGACACAGGCTTCGGGGCTGGCG	660
Db	602	TGCGCGCCGCTGTCTACACCATGAAACCTCTCTCTGTGACACAGGCTTCGGGGCTGGCG	661
QY	661	TGTTTCTCTGTGTACAGCTGGCTTGGGGCTTCAATTGTGATGTGACCTTAAGAGTGTAT	720
Db	662	TGTTTCTCTGTGTACAGCTGGCTTGGGGCTTCAATTGTGATGTGACCTTAAGAGTGTAT	721
QY	721	AAGTATGCGCTCTGGGGAGATACGCTTCACTGTGGGACCTTCTGGGTGTGGAGCG	780
Db	722	AAGTATGCGCTCTGGGGAGATACGCTTCACTGTGGGACCTTCTGGGTGTGGAGCG	781
QY	781	CTGCGCACTTACCGGGCGCGATGGCTTACCTCTCTGTAGAAAGATGGTTCGAAGAA	840
Db	782	CTGCGCACTTACCGGGCGCGATGGCTTACCTCTCTGTAGAAAGATGGTTCGAAGAA	841
QY	841	CTGTGCTTCCCGGTGTGGTCCAGAGAGGCGCGGTAGATGCACACCTGTGGCCACTGGAG	900
Db	842	CTGTGCTTCCCGGTGTGGTCCAGAGAGGCGCGGTAGATGCACACCTGTGGCCACTGGAG	901
QY	901	GAGCAGGTGCTTCTCACTGTGCAAGTGTGTGCCGACAGAGACTTTGGTGTAAATGTCGGCA	960
Db	902	GAGCAGGTGCTTCTCACTGTGCAAGTGTGTGCCGACAGAGACTTTGGTGTAAATGTCGGCA	961

QY	1261	AGGGGTTCGGTGGAGCCCCCGCCAGCTGGAGAGCCCGCAGCAATGTCACACCGCACAAG	1320
Db	1262	AGGGGTGGGTGGAGCCCCCGCCAGCTGGAGAGCCCGCAGCAATGTCACACCGCACAAG	1321
QY	1321	CCCTTATGACCCCTGGGCGCGGCTGGCTAGTGTACTTTCAGAGACCTTCTCCTT	1380
Db	1322	CCCTTATGACCCCTGGGCGCGGCTGGCTAGTGTACTTTCAGAGACCTTCTCCTT	1381
QY	1381	CCCTAGGGGCTGCAAGGGCCCTGTCCACAGCTCCTGTGGGGGTGGAGAGACTCCCTGGAGAG	1440
Db	1382	CCCTAGGGGCTGCAAGGGCCCTGTCCACAGCTCCTGTGGGGGTGGAGAGACTCCCTGGAGAG	1441
QY	1441	AGGCTGAGAAAGTGGAGAGGCTATGCTTTGGGGGAGCAGCCAGATGATGAAGTCTGGTCA	1500
Db	1442	AGGCTGAGAAAGTGGAGAGGCTATGCTTTGGGGGAGCAGCCAGATGATGAAGTCTGGTCA	1501
QY	1501	GAGCCAGAGTGGCTGGGCGCCAGCTGCTATGTATGAAGCTTCTTACTTTGTTCTAGACCCC	1560
Db	1502	GAGCCAGAGTGGCTGGGCGCCAGCTGCTATGTATGAAGCTTCTTACTTTGTTCTAGACCCC	1561
QY	1561	CACCCAGAGAACCAATCCAAATCAATGAAGTACATTCCAA	1599
Db	1562	CACCCAGAGAACCAATCCAAATGAAGTACATTCCAA	1600
		RESULT 2	
		US-10-023-282-90	
		Sequence 90, Application US/10023282	
		Publication No. US20030092893A1	
		GENERAL INFORMATION:	
		APPLICANT: Young et al.	
		TITLE OF INVENTION: 207 Human Secreted Proteins	
		FILE REFERENCE: P2007P1	
		CURRENT APPLICATION NUMBER: US/10/023, 282	
		CURRENT FILING DATE: 2001-12-20	
		EARLIER APPLICATION NUMBER: 09/205, 258	
		EARLIER FILING DATE: 1998-12-04	
		EARLIER APPLICATION NUMBER: PCT/US98/11422	
		EARLIER FILING DATE: 1998-06-04	
		EARLIER APPLICATION NUMBER: 60/048, 885	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/049, 375	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 881	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 880	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 896	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/049, 020	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 876	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 895	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 884	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 894	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 971	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 964	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 882	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 899	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 893	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 900	
		EARLIER FILING DATE: 1997-06-06	
		EARLIER APPLICATION NUMBER: 60/048, 901	
		EARLIER FILING DATE: 1997-06-06	

```

1  EARLIER APPLICATION NUMBER: 60/048, 892
2  EARLIER FILING DATE: 1997-06-06
3  EARLIER APPLICATION NUMBER: 60/048, 915
4  EARLIER FILING DATE: 1997-06-06
5  EARLIER APPLICATION NUMBER: 60/049, 019
6  EARLIER FILING DATE: 1997-06-06
7  EARLIER APPLICATION NUMBER: 60/048, 970
8  EARLIER FILING DATE: 1997-06-06
9  EARLIER APPLICATION NUMBER: 60/048, 972
10 EARLIER FILING DATE: 1997-06-06
11 EARLIER APPLICATION NUMBER: 60/048, 916
12 EARLIER FILING DATE: 1997-06-06
13 EARLIER APPLICATION NUMBER: 60/049, 373
14 EARLIER FILING DATE: 1997-06-06
15 EARLIER APPLICATION NUMBER: 60/048, 875
16 EARLIER FILING DATE: 1997-06-06
17 EARLIER APPLICATION NUMBER: 60/049, 374
18 EARLIER FILING DATE: 1997-06-06
19 EARLIER APPLICATION NUMBER: 60/048, 917
20 EARLIER FILING DATE: 1997-06-06
21 EARLIER APPLICATION NUMBER: 60/048, 949
22 EARLIER FILING DATE: 1997-06-06
23 EARLIER APPLICATION NUMBER: 60/048, 974
24 EARLIER FILING DATE: 1997-06-06
25 EARLIER APPLICATION NUMBER: 60/048, 883
26 EARLIER FILING DATE: 1997-06-06
27 EARLIER APPLICATION NUMBER: 60/048, 897
28 EARLIER FILING DATE: 1997-06-06
29 EARLIER APPLICATION NUMBER: 60/048, 898
30 EARLIER FILING DATE: 1997-06-06
31 EARLIER APPLICATION NUMBER: 60/048, 962
32 EARLIER FILING DATE: 1997-06-06
33 EARLIER APPLICATION NUMBER: 60/048, 963
34 EARLIER FILING DATE: 1997-06-06
35 EARLIER APPLICATION NUMBER: 60/048, 877
36 EARLIER FILING DATE: 1997-06-06
37 EARLIER APPLICATION NUMBER: 60/048, 878
38 EARLIER FILING DATE: 1997-06-06
39 EARLIER APPLICATION NUMBER: 60/070, 923
40 EARLIER FILING DATE: 1997-12-18
41 EARLIER APPLICATION NUMBER: 60/092, 921
42 EARLIER FILING DATE: 1998-07-15
43 EARLIER APPLICATION NUMBER: 60/094, 657
44 EARLIER FILING DATE: 1998-07-30
45 NUMBER OF SEQ ID NOS: 1227
46 SOFTWARE: PatentIn Ver. 2.0
47 SEQ ID NO 90
48 LENGTH: 1533
49 TYPE: DNA
50 ORGANISM: Homo sapiens
51 FEATURE:
52 NAME/KEY: SITE
53 LOCATION: (12)
54 OTHER INFORMATION: n equals a,t,g, or c
55 FEATURE:
56 NAME/KEY: SITE
57 LOCATION: (123)
58 OTHER INFORMATION: n equals a,t,g, or c
59 FEATURE:
60 NAME/KEY: SITE
61 LOCATION: (1522)
62 OTHER INFORMATION: n equals a,t,g, or c
63 FEATURE:
64 NAME/KEY: SITE
65 LOCATION: (1527)
66 OTHER INFORMATION: n equals a,t,g, or c
67 OTHER INFORMATION: n equals a,t,g, or c
68 OS=10-023-282-90

```

Query Match	89.2%	Score 1426.8	DB 9	Length 1533
Best Local Similarity	97.8%	Pred. NO. 0		
Matches 1459	Conservative 6	Mismatches 25	Indels 2	Gaps 2

109 GCCGCCACGGGCGACGGCCCCCAGCCAGCCAGGACCCCTTGGCAGCGGGAACCGCGGCTC 168

[illegible]

PRIOR APPLICATION NUMBER: 60/182,360
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: 60/191,261
PRIOR FILING DATE: 2000-03-22
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 3
LENGTH: 1759
TYPE: DNA
ORGANISM: Mus musculus
US-09-784-810A-3

Query Match 52.3%; Score 836.4; DB 10; Length 1759;
Best Local Similarity 72.8%; Pred. No. 7.9e-217;
Matches 1154; Conservative 0; Mismatches 411; Indels 21; Gaps 5;

QY 30 GGGGGGGAAGGCGAGCCCAACAGCCGCGCTGGAGCCCGCTGGGAGACCGATTA 89
DB 179 GAGGTGAGAGGGGAGCCCAAGCCAGTGGCCAGACACCTCTGGGCAACACGATTA 238
QY 90 GGAAGCTAAGGAGGAGCCCGG--CCAGGGGAGGCGCCCAACAGGCGGAGCC 146
DB 239 GAACTCAAGCAGAGGAGCCCGCTTACCTTACGAGCGCGGAGGAGCGGAGGCGCC 298
QY 147 CTGGCAGGAGGAGCGGCGGTGAGGTATGATCAAGCGGCGGCGCGGCGGCGT 206
DB 299 TTGTACAGGAGGAGCGGCGGTGAGGTATGATCAAGCGGCGGCGGCGGCGT 358
QY 207 CCCGCGGCGCTGCGGCGGTGAGGTATGATCAAGCGGCGGCGGCGGCGGCGT 266
DB 359 CCCGCGGCGCTGAGGTGAGGTATGATCAAGCGGCGGCGGCGGCGGCGT 418
QY 267 GCACCTCTCCGAGTACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 326
DB 419 GCACCTCTCCGAGTACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 478
QY 327 GATGCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 386
DB 479 GATGCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 538
QY 387 CTGGAGCGCTGCTGCTGATGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 446
DB 539 CTGGAGCGCTGCTGCTGATGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 598
QY 447 CATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 506
DB 599 AATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658
QY 507 CTCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 566
DB 659 CTCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 718
QY 567 TGAAGACCTCTGACCACTGACGCTATGCTGCTGCGCGGCTGCTGACCCATGAA 626
DB 719 TGAAGACCTCTGACCACTGACGCTATGCTGCTGCGCGGCTGCTGACCCATGAA 778
QY 627 CCTGCTGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 686
DB 779 CCTGCTGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 838
QY 687 GGGCTTATGCTGATGAGCACTAGAGAGGATGATGAGGAGGAGGAGGAGGAGG 746
DB 839 GGGCTTATGCTGATGAGCACTAGAGAGGATGATGAGGAGGAGGAGGAGGAGG 898
QY 747 CTTACTCTGGGCACTTCTGCTGCTGAGCGGCGGCGGAGCACTACCGCGGAGCTGCG 806
DB 899 TTTACAGTGGGCACTTCTTCTGCTGCTGAGCGGCGGAGCACTACCGCGGAGCTGCG 958
QY 807 TACCTCCCTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 866
DB 959 CTACCTTCTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1015
QY 867 GGGCGGCTAGATGACACACTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 926

DB 1016 GGGCGGCTGAGACACACCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1075
QY 927 GGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 986
DB 1076 GGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1135
QY 987 GATGCTTGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1046
DB 1136 GCTGTTGACAGACACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1195
QY 1047 GGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1106
DB 1196 TGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1255
QY 1107 TATGAGTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1166
DB 1256 TATGAGTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1315
QY 1167 CAAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1226
DB 1316 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1375
QY 1227 GGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1286
DB 1376 GGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1435
QY 1287 CTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1343
DB 1436 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1495
QY 1344 TGTGCTTATGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1403
DB 1496 TGTGCTTATGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1555
QY 1404 ACAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1463
DB 1556 AAGTCTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1613
QY 1464 CTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1513
DB 1614 ACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1673
QY 1514 TGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1573
DB 1674 CTGAGCACTGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1733
QY 1574 AATTCAAATTAAGTATGATTCCTCA 1599
DB 1734 AATTCAAATTAAGTATGATTCCTCA 1759

RESULT 5
US-09-970-516-5
Sequence 5, Application US/09970516
Patent No. US20020099029A1
GENERAL INFORMATION:
APPLICANT: NO. US20020099029A1artis AG
TITLE OF INVENTION: Induction of blood vessel formation through administration of
FILE REFERENCE: 4-31617
CURRENT APPLICATION NUMBER: US/09/970,516
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent In version 3.1
SEQ ID NO 5
LENGTH: 1149
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1149)
OTHER INFORMATION:

RESULT 10

US-09-817-676A-11
Sequence 11, Application US/09817676A
Patent No. US2002042101A1
GENERAL INFORMATION:
APPLICANT: Spiegel, Sarah
APPLICANT: Kohama, Takafumi
TITLE OF INVENTION: Mammalian Sphingosine Kinase Type 2 Isoforms, Cloning,
TITLE OF INVENTION: Expression and Methods of Use Thereof
FILE REFERENCE: 00170/HG
CURRENT APPLICATION NUMBER: US/09/817,676A
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/194,318
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 2698
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (387)..(2237)
PUBLICATION INFORMATION:
TITLE: Molecular cloning and functional characterization of a
TITLE: novel mammalian sphingosine kinase type 2 isoform
JOURNAL: J. Biol. Chem.
VOLUME: 275
ISSUE: 26
PAGES: 19513-19520
DATE: 2000-06-30
DATABASE ACCESSION NUMBER: AF245448
DATABASE ENTRY DATE: 2000-06-27
US-09-817-676A-11

Query Match 15.1% Score 242; DB 10; Length 2698;
Best Local Similarity 61.7%; Pred. No. 7.5e-56;
Matches 386; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 192 CCCCCGGGGGCTGCTCCCGGCGCTGCGCGCTGCTGCTGCTGAACCCGCGGCGG 251
DB 797 CACCCCTGAAATGCTGCCCGGAGAGCCAGCGCTGCTCATATGTCATATCCCTTTGGGG 856
QY 252 CAAGGCGAAGCGCTTGCAGCTTCCGAGAGTACGTGACGAGCCCTTTTGGCTGAGGCTGA 311
DB 857 GCGGGGCGCTGCGCTGCGCTGATGACCAAGCGGCTGATGATCTGGAAGCTGG 916
QY 312 AATCTCTTCACTGATGCTGATGAGCGGCGGGAACCAAGCGGCGGAGCTGTGGGTC 371
DB 917 GCTGCTCTTCACTGATGAGCGGCGGGAACCAAGCGGCGGAGCTGTGGGTC 976
QY 372 GGAGGAGCTGGCGCGCTGCGAGCTGTGCTGCTGCTGAGAGCGGCGCTGATGACGA 431
DB 977 GTTAAGCCTGATGATGAGGAGGATGCTGCTGCTGAGAGCGGCGCTGCTTACGA 1036
QY 432 GGTGTAAGCGGCTGATGAGCGGCGCTGAGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 491
DB 1037 GGTGCTGATGAGCGGCTGCTGATGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 1096
QY 492 TAGCTCCAGAGCGCTGCGAGCGGCGCTGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGG 551
DB 1097 TGTCTCCCTGATGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGG 1156
QY 552 TGAGAGCTGACCAATGAGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 611
DB 1157 TGAGAGCGGCTGCGGAGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 1216
QY 612 GCTGTAAGCGGCGGCTGCTGCTGCTGAGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 671
DB 1217 CAGCCTCTCTGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 1276
QY 672 GCTCAGCGGCGGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGG 731

DB 1277 CTTGCTGAGCGGCGGAGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 1336
QY 732 TCTGAGGAGAGCGGCTTACTGCTGAGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 791
DB 1337 CTTGAGGAGAGCGGCTTACTGCTGAGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 1396
QY 792 CCGGCGGAGCGGCTTACTGCTGAGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 817
DB 1397 CCGTGAAGCGGCTTACTGCTGAGCGGAGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGCTGTG 1422

RESULT 11

US-09-796-692-2905
Sequence 2905, Application US/09796692
Patent No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2905
TYPE: DNA
LENGTH: 199
ORGANISM: Homo sapiens
US-09-796-692-2905

Query Match 12.3% Score 197.4; DB 9; Length 199;
Best Local Similarity 99.5%; Pred. No. 6e-44;
Matches 198; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1088 CAGTGAAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1147
DB 1 CAGTGAAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 60
QY 1148 TCGGCTTCCGCTTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1207
DB 61 TCGGCTTCCGCTTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 120
QY 1208 TGGTGAAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1267
DB 121 TGGTGAAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 180

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 13:56:17 ; Search time 2385 Seconds

(without alignments)
10858.106 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600

Perfect score: 1599

Sequence: 1 cccccgggggctctctatagc.....aataaagtgcattcccaa 1599

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1558.6	97.5	2133	11	BC014439 Homo sapi
2	847.6	53.0	1054	14	BM808698 AGENCOURT
3	815.4	51.0	1137	14	BM810136 AGENCOURT
4	784.2	45.0	892	13	B1860351 AGENCOURT
5	757.4	47.4	888	14	B0647377 AGENCOURT
6	746.6	46.7	772	9	AT769914 wj30406.x

Result No.	Score	Query Match	Length	DB ID	Description
7	704.4	44.1	785	12	B680521 602628344
8	687.4	43.0	1100	10	BM916526 AGENCOURT
9	641.4	40.1	676	10	AM963415
10	635.6	39.7	702	10	BE464487
11	624.6	39.1	638	9	A1972156
12	615.8	38.5	662	13	BM386957
13	614.2	38.4	1005	13	BM559257 AGENCOURT
14	608.4	38.0	809	12	BC280830
15	606.8	37.9	1075	13	BM557357
16	605	37.8	793	12	BE678689
17	604.4	37.8	852	13	BE274434
18	600.2	37.5	875	13	BM006005
19	597	37.3	618	10	AM662445
20	582.4	36.4	971	10	AV694791
21	574.8	35.9	999	14	B0675531
22	572.6	35.8	603	10	BE275818
23	571.4	35.7	659	12	BE891653
24	569.6	35.6	949	12	BC026631
25	566.6	35.4	606	14	BQ251614
26	527	33.0	720	14	BQ109456
27	524.6	32.8	947	13	B1415545
28	518	32.4	856	13	B1757316
29	515.2	32.2	2145	11	BC011432
30	507	31.7	842	13	B1648186
31	505.2	31.6	587	9	AU147968
32	504.8	31.6	1134	14	BM808666
33	491	30.7	1115	13	BM550039
34	486	30.4	697	13	B1255900
35	476.2	29.8	534	12	BE904632
36	467.4	29.2	914	13	B1823329
37	457	28.6	732	12	BE728564
38	455.2	28.5	478	12	BE740866
39	446.8	27.9	1232	13	BM556915
40	440	27.5	447	9	A1042283
41	421.2	26.3	650	12	BG104264
42	418	26.1	675	9	AU134361
43	414.2	25.9	733	14	BM934157
44	411.8	25.8	973	10	BE273426
45	405	25.3	502	10	BE236257

ALIGNMENTS

RESULT 1
LOCUS BC014439 2133 bp mRNA linear HTC 19-DEC-2001
DEFINITION Homo sapiens, clone IMAGE:4871343, mRNA.
ACCESSION BC014439
VERSION BC014439.1 GI:17939597
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2133)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DT/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
Info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Gulin,

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhun, Parvaneh Seedy, Jacqueline Schein, Duane Smallue, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Nastasia van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
Series: IRAL Plate: 34 Row: 1 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 11464966
This clone has the following problem: incomplete processing.

FEATURES

source

Location/Qualifiers
1. 2133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4871343"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="DH10B-R"
/lab_host="NIH-MGC_49"
/note="Vector: pOB7"

BASE COUNT

350 a 681 c 719 g 383 t

ORIGIN

Query Match 97.5%; Score 1558.6; DB 11; Length 2133;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1573; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

OY 3 CCCCAGGGCTCTTATAGCCAGGCTCCGGGGGAGAGAGCCGACGCGGCTG 62
    |||||
DB 521 CCGCGGTGCTCTACAGCCAGGCTCCGGGGGAGAGAGCCGACGCGGCTG 580
OY 63 CAGACGCCCTGGGAGAGCAGATAGAGCTGAAGCAGAGAGCCGCGCAGCGG 122
    |||||
DB 581 CAGAGCCCGCTGGAGAGCAGATAGAGCTGAAGCAGAGAGCCGCGCAGCGG 640
OY 123 CCGCCCGACAGCGCAGAGAGCCCTGCGAGCGGAGAGCCGCGGCTGATGATCC 182
    |||||
DB 641 CCGCCCGACAGCGCAGAGAGCCCTGCGAGCGGAGAGCCGCGGCTGATGATCC 700
OY 183 AGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 242
    |||||
DB 701 AGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 760
OY 243 GCGGGGGGGAAGGGGAGGCTGTCAGCTCTTCGAGAGTCAGAGCCGCGGAGCT 302
    |||||
DB 761 GCGGGGGGGAAGGGGAGGCTGTCAGCTCTTCGAGAGTCAGAGCCGCGGAGCT 820
OY 303 TGAGGCTGAATTCCTTACAGGCTGATGCTCACTGAGCGGAGAGCCGCGGAGCT 362
    |||||
DB 821 TGAGGCTGAATTCCTTACAGGCTGATGCTCACTGAGCGGAGAGCCGCGGAGCT 880
OY 363 GGTGGCTGAGAGAGAGTGGCGGCTGAGAGCGCTCTGTGATGCTGAGAGCGGCT 422
    |||||
DB 881 GGTGGCTGAGAGAGAGTGGCGGCTGAGAGCGCTCTGTGATGCTGAGAGCGGCT 940
OY 423 GATGACAGAGAGTGGAGAGGCTCATGAGAGCGGCTGATGAGAGAGCCGAGTCA 482
    |||||
DB 941 GATGACAGAGAGTGGAGAGGCTCATGAGAGCGGCTGATGAGAGAGCCGAGTCA 1000
OY 483 GCGCCCTGTAGCTCCAGCAGAGGCTTGAGAGCGGCTGAGAGCTCTTGAACATTA 542
    |||||
DB 1001 GCGCCCTGTAGCTCCAGCAGAGGCTTGAGAGCGGCTGAGAGCTCTTGAACATTA 1060
OY 543 TGTGCTGTAGAGAGAGTGAACATGAGAGCTCTGAGAGCTGAGAGCTGATGCTG 602
    |||||
DB 1061 TGTGCTGTAGAGAGAGTGAACATGAGAGCTCTGAGAGCTGAGAGCTGATGCTG 1120
OY 603 CCGCCCGAGTGTGACAGCATGAACCTGCTGTCTGTGACAGAGGCTTGAGAGCTG 662
    |||||
DB 1121 CCGCCCGAGTGTGACAGCATGAACCTGCTGTCTGTGACAGAGGCTTGAGAGCTG 1180
```

```

OY 663 GTTCTGTGTGACAGCTTGAGCTGGGGCTTCAATTCGATGAGAGCTGAGATATA 722
    |||||
DB 1181 CTCTCTGTGTGACAGCTTGAGCTGGGGCTTCAATTCGATGAGAGCTGAGATATA 1240
OY 723 GTTCTGTGTGACAGCTTGAGCTGGGGCTTCAATTCGATGAGAGCTGAGATATA 782
    |||||
DB 1241 GTTCTGTGTGACAGCTTGAGCTGGGGCTTCAATTCGATGAGAGCTGAGATATA 1300
OY 783 GCGACCTTACAGCTGGGGCTTCAATTCGATGAGAGCTGAGATATA 842
    |||||
DB 1301 GCGACCTTACAGCTGGGGCTTCAATTCGATGAGAGCTGAGATATA 1360
OY 843 TGTCTTCCCGCTTGTGTGAGAGAGGCGCGGAGATGAGAGCTGAGAGCTGAGAG 902
    |||||
DB 1361 TGTCTTCCCGCTTGTGTGAGAGAGGCGCGGAGATGAGAGCTGAGAGCTGAGAG 1420
OY 903 GAGAGCTTCTTCACTGAGAGTGTGAGAGAGGAGAGCTTGTGAGAGCTGAGAGCT 962
    |||||
DB 1421 GCGAGTGGCTTCTTCACTGAGAGTGTGAGAGAGGAGAGCTTGTGAGAGCTGAGAG 1480
OY 963 GGTGACATGAGAGCTGAGAGTGTGAGAGAGGAGAGCTTGTGAGAGCTGAGAGCT 1022
    |||||
DB 1481 GGTGACATGAGAGCTGAGAGTGTGAGAGAGGAGAGCTTGTGAGAGCTGAGAGCT 1540
OY 1023 GGTGACATGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCTGAGAGCTTGT 1082
    |||||
DB 1541 GGTGACATGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCTGAGAGCTTGT 1600
OY 1083 CCTGACATGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCTTGT 1142
    |||||
DB 1601 CCTGACATGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCTTGT 1660
OY 1143 CGTGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1202
    |||||
DB 1661 CGTGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1720
OY 1203 ATTGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1262
    |||||
DB 1721 ATTGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1780
OY 1263 CGTGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1322
    |||||
DB 1781 CGTGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1840
OY 1323 CTTATGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1382
    |||||
DB 1841 CTTATGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1900
OY 1383 CTTATGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1442
    |||||
DB 1901 CTTATGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1960
OY 1443 GGTGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1502
    |||||
DB 1961 GGTGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 2020
OY 1503 GCGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 1562
    |||||
DB 2021 GCGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGCT 2080
OY 1563 CCGCAGAGAGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCT 1599
    |||||
DB 2081 CCGCAGAGAGAGAGCTTGTGAGAGAGGAGAGCTTGTGAGAGAGGAGAGCT 2117
```

RESULT 2
BM808698 1054 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6582622 NIH_MGC_41 Homo sapiens cdna clone IMAGE:5471258
DEFINITION 5', mRNA sequence.
ACCESSION BM808698
VERSION BM808698.1 GI:19125509
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mhc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1979 row: 9 column: 03
High quality sequence start: 7
High quality sequence stop: 666.
Location/Qualifiers

FEATURES
SOURCE
1. 1054
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5471258"
/clone.lib="NIH-MGC 41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site:1: XhoI, Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
SuperScript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 177 a 317 c 340 g 220 t
ORIGIN

Query Match 53.0%; Score 847.6; DB 14; Length 1054;
Best Local Similarity 92.1%; Pred. No. 3.4e-176;
Matches 962; Conservative 0; Mismatches 69; Indels 13; Gaps 6;

```

OY 317 CTTACGCTGATGCTACTGAGCGGGAACCAACGCGGGAGTGTGGCGGTGGAG 376
    |||||||
DB 1 CTTACGCTGATGCTACTGAGCGGGAACCAACGCGGGAGTGTGGCGGTGGAG 60
OY 377 AGCTGGGCGGTGGAGCGCTGTGTGTCATGTCGAGAGCGGCTGATGACGAGTGG 436
    |||||||
DB 61 AGCTGGGCGGTGGAGCGCTGTGTGTCATGTCGAGAGCGGCTGATGACGAGTGG 120
OY 437 TGAACGGGCTATGAGAGCGGCTGTGACGAGACCGCATCCAGAACCCCTGTATGCC 496
    |||||||
DB 121 TGAACGGGCTATGAGAGCGGCTGTGACGAGACCGCATCCAGAACCCCTGTATGCC 180
OY 497 TCCGAGGAGGCTGTGCAAGCGGCTGGAGGCTCTTAAACATATATGCTGTATGAGC 556
    |||||||
DB 181 TCCGAGGAGGCTGTGCAAGCGGCTGGAGGCTCTTAAACATATATGCTGTATGAGC 240
OY 557 AGGTACCAATGAAGACCTCTGACCAATGACACCTATTTGCTGTGCCCGGCTGTGT 616
    |||||||
DB 241 AGGTACCAATGAAGACCTCTGACCAATGACACCTATTTGCTGTGCCCGGCTGTGT 300
OY 617 CACCATGATACCTGCTGTCTGTGACACAGCGCTGGGCGCTGCTTCTGTGTCTCA 676
    |||||||
DB 301 CACCATGATACCTGCTGTCTGTGACACAGCGCTGGGCGCTGCTTCTGTGTCTCA 360
OY 677 GCGTGGCGTGGGCGCTTCAATGCTGTGATGTGAGACTGAGATGATAGTGGCGCTGG 736
    |||||||
DB 361 GCGTGGCGTGGGCGCTTCAATGCTGTGATGTGAGACTGAGATGATAGTGGCGCTGG 420
OY 737 GGGAGATGCGCTTCACTGTGGGCACTTCTGCGCTGTGAGCGCCCTGCGACATACCGG 796
    |||||||

```

DB 421 GGGAGATGCGCTTCACTGTGGGCACTTCTGCGCTGTGAGAGCCCTGGCACATACCGG 480
OY 797 GCCAGCTGCTACCTCCTCTAGAGAGTGGTTTAAACACACTCTCTCCCGCTTG 856
DB 481 GCCAGCTGCTACCTCCTCTAGAGAGTGGTTTAAACACACTCTCTCCCGCTTG 540
OY 857 TGGTCCACAGAGGCGCGGTAGATGACACACTGTTGCCACTGAGAGAGGAGGCGCTTG 916
DB 541 TGGTCCACAGAGGCGCGGTAGATGACACACTGTTGCCACTGAGAGAGGAGGCGCTTG 600
OY 917 ACTGGCAGTGTGCGCCAGCAGAGACTTGTGCTAGTCTGACACTGCTGACACTGAC 976
DB 601 ACTGGCAGTGTGCGCCAGCAGAGACTTGTGCTAGTCTGACACTGCTGACACTGAC 660
OY 977 TGGCCAGTGAATGTTTCTGACACCATGGGCGCGTGTGACAGTGGCTCATGTCTGT 1036
DB 661 TGGCCAGTGAATGTTTCTGACACCATGGGCGCGTGTGACAGTGGCTCATGTCTGT 720
OY 1037 TCTACGTGC-GAGGCGAGTGTCTGTCGTCATGCTGCTGCGCTCTTCTGGCCATGAG 1095
DB 721 TCTACGTGCAGGCGGCGGAGTGTCTGTCGTCATGCTGCTGCGCTCTTCTGGCCATGAG 780
OY 1096 AAGGCGAGCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1155
DB 781 AAGGCGAGCATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
OY 1156 CGCTTGGAGCCCAAGAGATGGGAAAGTGTGTTTGAAGTGA--TGCGAATTTGATGTTA 1213
DB 841 CGCTTGGAGCCCAAGAGATGGGAAAGTGTGTTTGAAGTGAAGTGAAGTGAAGTGA 900
OY 1214 GC--GAGGCGCGGAGGCGGAGTGTGACCACTTCTGATGATGATGATGATGATGATG 1267
DB 901 GCGGAGGCGCGGAGGCGGAGTGTGACCACTTCTGATGATGATGATGATGATGATG 960
OY 1268 GCGTGGAGCCCGCGCGCC--ACCTGGAACCCCGACAGATGCC--ACCGCAGAAAGGCC 1323
DB 961 GCGTGGAGCCCGCGCGCCCGCGCGAGTGAACCCCGCGAGATGCCACCGCAGAAAGGCC 1020
OY 1324 TTATGACCCCTGTGGCGCGCTGTG 1347
DB 1021 TTTTATCCCTGTGGCGCGCGCTGTG 1044

RESULT 3
BM810136 1137 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6579661 NIH-MGC_41 Homo sapiens cDNA clone IMAGE:5470089
DEFINITION 5', mRNA sequence.
ACCESSION BM810136
VERSION BM810136.1 GI:19126959
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mhc.nci.nih.gov/
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1976 row: f column: 10
High quality sequence stop: 623.
Location/Qualifiers
1. 1137
/organism="Homo sapiens"

Query Match	51.0%	Score 815.4	DB 14	Length 1137
Best Local Similarity	94.2%	Pred: No.1-Le-169		
Matches 888	Conservative 0	Mismatches 47	Indels 8	Gaps 4
QY	317	CCTTCACGCTGATGCTCTACTGAGCGCGGACACCCAGCGCGGAGAGCTGGTGGGTCGAGG	376	
Db	1	CCCTTCACGCTGATGCTCTACTGAGCGCGGACACCCAGCGCGGAGAGCTGGTGGGTCGAGG	60	
QY	377	AGCGGGGCGCTGTGGAGCGCTCTGTGTATGTCTGTGAGACGGGCTGTATGACAGAGGTGG	436	
Db	61	AGCTGGGCGCGCTGTGGAGCGCTCTGTGTATGTCTGTGAGACGGGCTGTATGACAGAGGTGG	120	
QY	437	TGAACGGGCTCATGAGACGGCGCTGACTGGGAGACCGGCATTCAGAAAGCCCTGTATACC	496	
Db	121	TGAACGGGCTCATGAGACGGCGCTGACTGTGAGAGACCGGCATTCAGAAAGCCCTGTGTAGCC	180	
QY	497	TCCGACGAGGCTCTGGCAACGCGGTGGACGCTTCTTGAACCATTTATGCTGGCTATAGGC	556	
Db	181	TCCGACGAGGCTCTGGCAACGCGGTGGACGCTTCTTGAACCATTTATGCTGTATAGAC	240	
QY	557	AGGTACCAATGAAGAACCTCTGTGACCAATGACAGCATATGTGCTGTACCGCCGCTGTCT	616	
Db	241	AGGTACCAATGAAGAACCTCTGTGACCAATGACAGCATATGTGCTGTACCGCGGCTGTCTGT	300	
QY	617	CACCCATGAACCTGTCTGTCTGTGACACAGGCTTCTGGGGCTTGGCGCTGGTCTCTGTGTCCTCA	676	
Db	301	CACCCATGAACCTGTGTCTGTCTGTGACACAGGCTTCTGGGGCTTGGCGCTGTCTGTGTCTCA	360	
QY	677	GCGTGGCGCTGGGGGCTTCAATTGCTCATGTGAGACCTAGAGAGTGAATAGTATCGGCTGTGG	736	
Db	361	GCGTGGCGCTGGGGGCTTCAATTGCTCATGTGAGACCTAGAGAGTGAATAGTATCGGCGTGTGG	420	
QY	737	GGAAGATGCGCTTCACTCTGTGGACCTTCTGCGCTCTGGGACAGCCGTCGGACCTACCGGG	796	
Db	421	GGAAGATGCGCTTCACTCTGTGGACACCTTCTGCGCTCTGGGACAGCCCTGCGACCTACCGGG	480	
QY	797	GCGCACTGGCTACCTCCCTGTAGAGAAAGTGGGTTTTCAGACACACTGCTTCCCGCGTG	856	
Db	481	GCGCACTGGCTTACCTCCCTGTAGAGAAAGTGGGTTTTCAGACACACTGCTTCCCGCGTG	540	
QY	857	TGGTCCAGACAGGCGCCCGGTAGATGACACACTGTGTGCCACTGTGAGAGAGAGGTGCTTCTC	916	
Db	541	TGGTCCAGACAGGCGCCCGGTAGATGACACACTGTGTGCCACTGTGAGAGAGAGGTGCTTCTC	600	
QY	917	ACTGGACAGGTGTGTCGCCGACGAGAGACTTGTGTCTAGTCTGTGGACACTGCTGCATCGCAC	976	
Db	601	ACTGGACAGGTGTGTCGCCGACGAGAGACTTGTGTCTAGTCTGTGGACACTGCTGCATCGCAC	660	
QY	977	TGGCCAGTGAAGATTGTTGCTGCACCCATGGGCGCGGTGACACTGGGCTCACTGATCATCTGT	1036	
Db	661	TGGCCAGTGAAGATTGTTGCTGCACCCATGGGCGCGGTGACACTGGGCTCACTGATCATCTGT	720	
QY	1037	TCATACGTGCGGGGAGTGTCTCATGCAATCTGTGCTGCGCTCTTCTGACCATGAGAGA	1096	
Db	721	TCATACGTGCGGGGAGTGTCTCATGCAATCTGTGCTGCGCTCTTCTGACCATGAGAGA	780	
QY	1097	AGGGCAGCATATGAGATATGAAATGCCCTACTTGGTATATATGG-CCCGTGGTCCCTTTC	1156	

RESULT 4	892 bp	mrna	linear	EST 10-OCT-2001
BT860351				
LOCUS	BT860351			
DEFINITION	6033874.971 NIH_MGC_87 Homo sapiens CDNA clone IMAGE:5396178 5',			
	mrna sequence.			

ORGANISM	REFERENCE
<i>Homo sapiens</i>	NIH-MGC http://mhc.nci.nih.gov/ , 1 (bases 1 to 892)
Euteleostomi; Euteleostomi; Euteleostomi; Metazoa; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	

Email: cgapops@leuven.be; DCPD/DPD
Tissue Procurement: DCPD/DPD
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information center
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.jov>
Plate: L14M1209 row: n column: 19
High quality sequence stop: 878.

FEATURES	Location/Qualifiers
SOURCE	1. .892
	2. .892- "Homo sapiens"

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5396178"
/clone_id="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

```

BASE COUNT	152 a	266 c	483 g	191 t
ORIGIN				
Query Match		49.08;	Score 784.2;	DB 13; Length 892;
Best Local Similarity	95.2%;		Pred. No. 2.9e-162;	
Matches 852; Conservative	0;	Mismatches 38;	Indels 5;	Gaps 4

QY	428	ACGAGGTGTGTAACGGGCTCATGTAGACGGGCTTACGTGGAGACGGGCATTCAGAAACCCC	487
Db	1	ACGAGGTGTGTAACGGGCTCATGTAGAGGGGCTTGTCTGGAGACCGCCATCCAGAAACCCC	60
QY	488	TGTGTAGCTCTCCAGACGAGCTCTGTGGCAACGGCGTGGGCACTCTCTTGAAACATTATGCTG	547
QY	61	TGTGTAGCTCTCCAGACGAGCTCTGTGGCAACGGCGTGGGCACTCTCTTGAAACATTATGCTG	120
Db	548	GCTATGACAGGTCACCAATGAAGACTCTCCGACCAACGTACGAGCTATTCTGTGGCGGC	607
QY	121	GCTATGACAGGTCACCAATGAAGACTCTCCGACCAACGTACGAGCTATTCTGTGGCGGC	180
QY	608	CGGTCTGTGCACCATGAACTGTGCTTGTGCACACAGGCTTGGGGCTGGGCTCTTCT	667

Db	181	GGCGTGTGTACACCATGAACTGTGTGTCTGTGCACAGGGCTTGTGGGGCTGTGGCTCTTCT	240
OY	668	CTGTGTCTACGCTTGCGCTGGGGCTTCATTTGTGTATGTGTGACCTAGAGAAGTATGATTC	727
Db	241	CTGTGTCTACGCTTGCGCTGGGGCTTCATTTGTGTATGTGTGACCTAGAGAAGTATGATTC	300
OY	728	GGCGTGTGTGAGATGGGCTTCATCTGTGGGACCTTCCTGGCTGTGGACGCTTGCGCA	787
Db	301	GGCGTGTGTGAGATGGGCTTCATCTGTGGGACCTTCCTGGCTGTGGACGCTTGCGCA	360
OY	788	CCCTACCGCGCGCACTGGCTACCTCCCTGTAGAGAGATGGGCTTTCAGACACTTGCTT	847
Db	361	CCCTACCGCGCGCACTGGCTACCTCCCTGTAGAGAGATGGGCTTTCAGACACTTGCTT	420
OY	848	CCCCCGTTTGTGTCCAGCAGGGCCCCGGTAGATGCACACCTGTGTCCACTGTGAGAGCAGG	907
Db	421	CCCCCGTTTGTGTCCAGCAGGGCCCCGGTAGATGCACACCTGTGTCCACTGTGAGAGCAGG	480
OY	908	TGCGCTTCATCTGGCAGGTGTGTGCCCCGACGAGACCTTGTGTGTCTGTCTGTGCACTGTGC	967
Db	481	TGCGCTTCATCTGGCAGGTGTGTGCCCCGACGAGACCTTGTGTGTCTGTCTGTGCACTGTGC	540
OY	968	ACTGCACTTGCGCGCAGTGAATGTTTGTGTGACCATAGGCGCGTGTGACACTGGCTCA	1027
Db	541	ACTGCACTTGCGCGCAGTGAATGTTTGTGTGACCATAGGCGCGTGTGACACTGGCTCA	599
OY	1028	TGCATCTGTCTACGTGTGCGGGCGGAGTGTCTGTGCCATGTCTGTGCGCTCTTCTTG	1087
Db	600	TGCATCTGTCTACGTGTGCGGGCGGAGTGTCTGTGCCATGTCTGTGCGCTCTTCTTG	659
OY	1088	CCATGTGAGAAAGGCGAGCATATGTGAGATGAATCCCTACTTGTGTATGTGGCCGTG	1147
Db	660	ACATGTGAGAAAGGCGAGCATATGTGAGATGAATCCCTACTTGTGTATGTGGCCGTG	719
OY	1148	TCGCGCTTCCGCTTGTGGAGCCCAAGATG-CGAAAGTGTGTGTGCAAGTGTGAGGAAATTG	1206
Db	720	TCGCGCTTCCGCTTGTGGAGCCCAAGATGTGTGTGTGTGTGCAAGTGTGAGGAAATTG	779
OY	1207	ATGCTTACGAGGCCGTGTGACGCGCCAGGTGTCACCCAACTACTTGTGTGATGTGACGCGT	1266
Db	780	ATGCTTACGAGGCCGTGTGACGCGCCAGGTGTCACCCAA-FACTTCTGTGATGTGACGCGT	838
OY	1267	TGCGTGTGAGCCCCCGCCACGTGTGGAAGCCCCCAGACAGATGCCACCGCCAGAAAGAC	1321
Db	839	TGC--TGTGAGCCCCCGCCACGTGTGGAAGCCCCCAGACAGATGCCACCGCAAAAAAGC	891

RESULT 5
 B0647377
 LOCUS
 DEFINITION
 AGNCNCURJ_8414975 NIH_MGC_100 Homo sapiens CDNA clone IMAGE:52720711
 5', mRNA sequence.
 ACCESSION
 VERSION
 B0647377.1 GI:21771549
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CCAP (Stanford)
 cDNA Library Preparation: Rudin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2449 row: n column: 08

FEATURES		source
High quality sequence start: 6 High quality sequence stop: 668. Location/Qualifiers 1..888		
/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6272071" /clone_lib="NIH_MGC_100" /tissue_type="hepatocellular carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="organ: liver; Vector: pOMB7, site 1: XhoI; site 2: EcoRI; cDNA made by oligo-dT priming. Directionality cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGACGAG(g). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library." BASE COUNT 137 a 274 c 284 g 193 t ORIGIN		
Query Match	47.4%; Score 757.4; DB 14; Length 888;	
Best Local Similarity	94.9%; Pred. No. 2.3e-156;	
Matches 815; Conservative	0; Mismatches 41; Indels 3; Gaps 3	
582	CACATGCACGCTATTGCTGTGTGCGCGCGCGGTGCTGCACCATGAACCTGCTCTCTGCA	641
Db	30 CCACGTGACGAGTATTGCTGTGCGCGCGCGGTGCTGCACCATGAACCTGCTCTCTGCA	89
QY	642 CACGCGCTTCGGGGCTGCGCGCTCTCTCTGTGCTCAGCCTGGCGCTGGGCTTCATTGCTGA	701
Db	90 CACGCGCTTCGGGGCTGCGCGCTCTCTCTGTGCTCAGCCTGGCGCTGGGCTTCATTGCTGA	149
QY	702 TGTGACCTAGAGAGTGTAAATATATCAGGCGTGTGGGGAGATAGCCTTCACTCTGGGAC	761
Db	150 TGTGACCTAGAGAGTGTAAATATATCAGGCGTGTGGGGAGATAGCCTTCACTCTGGGAC	209
QY	762 CTTCCTGGGTGTGGACGCGCTGCGGACCTACCGCGGCGGAGTGGTACCTCCCTGTAGG	821
Db	210 CTTCCTGGGTGTGGACGCGCTGCGGACCTACCGCGGCGGAGTGGTACCTCCCTGTAGG	269
QY	822 AAGATGGGTTTCAAGACACTGCTTCCCGCTTGTGTGCTCAGCAGGCGCGGTATGATGC	881
Db	270 AAGATGGGTTTCAAGACACTGCTTCCCGCTTGTGTGCTCAGCAGGCGCGGTATGATGC	329
QY	882 ACACCTGGGGCCACAGGAGAGAGACAGATGCTCTCTACAGGAGGAGTGGCCGACGAGGA	941
Db	330 ACACCTGGGGCCACAGGAGAGAGACAGATGCTCTCTACAGGAGGAGTGGCCGACGAGGA	389
QY	942 CTTTGTGTAGTCTCGGCACTGCTCAGCTCGGACCTGGCCAGATGAGATGTTGTGACCC	1001
Db	350 CTTTGTGTAGTCTCGGCACTGCTCAGCTCGGACCTGGCCAGATGAGATGTTGTGACCC	449
QY	1002 CATGGGCGCTGTGACGCTGGCGTATGCAATGTTTCACTGTGGGCGGAGTGTCTCG	1061
Db	450 CATGGGCGCTGTGACGCTGGCGTATGCAATGTTTCACTGTGGGCGGAGTGTCTCG	509
QY	1062 TGGCATGCTGCTGCGCCCTTCTCTGCGCATGAGAGAGGACGAGATATGAGATGATG	1122
Db	510 TGGCATGCTGCTGCGCCCTTCTCTGCGCATGAGAGAGGACGAGATATGAGATGATG	569
QY	1122 CCCCACCTTGGATATGTGCCGTGTGGCCTTCCGCTTGGAGCCCAAGATGGGAAAG	1181
Db	570 CCCCACCTTGGATATGTGCCGTGTGGCCTTCCGCTTGGAGCCCAAGATGGGAAAG	629
QY	1182 TGTGTTTGCAGTGGAGGAAATGATGGTTAGCAGAGCCCTGCAGGGCCAGGTACACC	1241
Db	630 TGTGTTTGCAGTGGAGGAAATGATGGTTAGCAGAGCCCTGCAGGGCCAGGTACACC	689
QY	1242 AAACACTTCTTGATGTGTGTCAGC-GGTTGCTGTGAGACCCCGCCGAC-CTGGAAGCCGAC	1299
Db	690 AAACACTTCTTGATGTGTGTCAGC-GGTTGCTGTGAGACCCCGCCGACCTGGAAGCCGAC	749

http://image.llnl.gov
Plate: ILAMI0612 row: 9 column: 12
High quality sequence stop: 783.
Location/Qualifiers

FEATURES

1. 785

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4753187"
/clone_lib="NCI-CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPO6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI-CGAP library."

BASE COUNT 130 a 231 c 250 g 174 t

ORIGIN

Query Match 44.1%; Score 704.4; DB 12; Length 785;

Best Local Similarity 96.7%; Pred. No. 1e-144;
Matches 762; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

OY 545 CTGGCTATGACAGGTACCAATGAAGACTCTGACCACTGCACGCTATTGCTGCC 604
DB 1 CTGGCTATGACAGGTACCAATGAAGACTCTGACCACTGCACGCTATTGCTGCC 60
OY 605 GCCCGGTCTGTACACCATGAACCTGCTGTCTGTACACAGGCTTGGGGCTGGCTCT 664
DB 61 GCCCGGTCTGTACACCATGAACCTGCTGTCTGTACACAGGCTTGGGGCTGGCTCT 120
OY 665 TCTGTGTGTCAGCTGCTGGGCTGCTCATTTGCTGATGTGGACCTAGAGATGATAGT 724
DB 121 TCTGTGTGTCAGCTGCTGGGCTGCTCATTTGCTGATGTGGACCTAGAGATGATAGT 180
OY 725 ATCGGCTGTGGGGAGATGCGCTTCACTGTGGCACCTTCTGCTGTGGCACCTGTC 784
DB 181 ATCGGCTGTGGGGAGATGCGCTTCACTGTGGCACCTTCTGCTGTGGCACCTGTC 240
OY 785 GCACCTACCGGGGCTGCTGCTACCTCTCTGTGGAAGAGTGGGTTTCAAGAACCTG 844
DB 241 GCACCTACCGGGGCTGCTGCTACCTCTCTGTGGAAGAGTGGGTTTCAAGAACCTG 300
OY 845 CTTCCCGGCTGTGTGTCAGAGAGGCGGCTGATGATGACACCTGTGCTGCTGAGAGC 904
DB 301 CTTCCCGGCTGTGTGTCAGAGAGGCGGCTGATGATGACACCTGTGCTGCTGAGAGC 360
OY 905 AGGTGCTTGTCACTGTGAGGTGTGCGCGAGAGAGACTTGTGCTAGTCTGTGACCTG 964
DB 361 AGGTGCTTGTCACTGTGAGGTGTGCGCGAGAGAGACTTGTGCTAGTCTGTGACCTG 420
OY 965 TGCACCTGACCTGGCCAGTGAAGTGTGCTGACCCATGGGCGGCTGTGACCTGGCG 1024
DB 421 TGCACCTGACCTGGCCAGTGAAGTGTGCTGACCCATGGGCGGCTGTGACCTGGCG 480
OY 1025 TCATGCACTGTGTCAGTGGGCGGAGTGTCTGCTGCACTGCTGTGCGCTCTTCC 1084
DB 481 TCATGCACTGTGTCAGTGGGCGGAGTGTCTGCTGCACTGCTGTGCGCTCTTCC 540
OY 1085 TGGCCATGAGAGAGGAGCATATGAGATGAATGATGATGATGATGATGATGATG 1144
DB 541 TGGCCATGAGAGAGGAGCATATGAGATGAATGATGATGATGATGATGATGATG 600
OY 1145 TGGTGGCTTGGCGGCTGGAGAGGAGAGTGGAGAGTGGTGGAGTGGATGGAG 1203
DB 601 TGGTGGCTTGGCGGCTGGAGAGGAGAGTGGAGAGTGGTGGAGTGGATGGAG 660
OY 1204 TTGATGTTAGAGAGGCGCTGAGAGGCGAGGTGACACCAAACTACTTGTGATGTCAGC 1263
DB 661 TTGATGTTAGAGAGGCGCTGAGAGGCGAGGTGACACCAAACTACTTGTGATGTCAGC 718
OY 1264 GGTTCGTGAGAGCCCGCCCACTGTAAGAGCCCAAGAGATGACACCCCAAGAGAGCC 1323
DB 719 GGTTCGTGAGAGCCCGCCCACTGTAAGAGCCCAAGAGATGACACCCCAAGAGAGCC 776

OY 1324 TTATGACC 1331
DB 777 TTATGACC 784

RESULT 8
LOCUS BM916526 1100 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6641989 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482958
5', mRNA sequence.
ACCESSION BM916526
VERSION BM916526.1 GI:19366905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1100)
NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2009 row: n column: 15
High quality sequence stop: 625.

FEATURES

source

1. 1100

Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5482958"
/clone_lib="NIH-MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site: 1; XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 187 a 333 c 348 g 222 t 10 others
ORIGIN

Query Match 43.0%; Score 687.4; DB 14; Length 1100;

Best Local Similarity 87.9%; Pred. No. 5.9e-141;
Matches 830; Conservative 0; Mismatches 93; Indels 21; Gaps 7;

OY 317 CTTTACGCTGATGCTCACTGAGCGCGGAGCAACGCGCGAGCTGTGGGTTGGAG 376
DB 1 CTTTACGCTGATGCTCACTGAGCGCGGAGCAACGCGCGAGCTGTGGGTTGGAG 60
OY 377 AGCTGGGCGGCTGAGGAGCTGTGATGATGATGATGATGATGATGATGATGATG 436
DB 61 AGCTGGGCGGCTGAGGAGCTGTGATGATGATGATGATGATGATGATGATGATG 120
OY 437 TGAAGGCGCTATGAGAGGCGCTGATGAGAGGAGCGCATCCAGAGCCCTGTGTAGCC 496
DB 121 TGAAGGCGCTATGAGAGGCGCTGATGAGAGGAGCGCATCCAGAGCCCTGTGTAGCC 180
OY 497 TCCAGAGAGCTGTGAGAGGCGCTGAGAGGAGCTTCTTGAACATTAATGCTGATGAGC 556
DB 181 TCCAGAGAGCTGTGAGAGGCGCTGAGAGGAGCTTCTTGAACATTAATGCTGATGAGC 240
OY 557 AGGTACCAATGAAGACTCTGACCACTGACAGCTATTGCTGTGCGCGCGGTGTGT 616

```

|||||
Db 241 AGGTACCAATGAAAGACCTCTGACCAACCTGACACCTATGTTGCGCGCGGCTGCTGT
|||
Qy 617 CACCCATGAAACCTGCTGCTGACACAGGCTTGGGGCTGCGCTGCTCTGTGCTCA 676
|||
Db 301 CACCCATGAAACCTGCTGCTGCTGACACAGGCTTGGGGCTGCGCTGCTCTGTGCTCA 360
|||
Qy 677 GCGTGGCTGGGGCTTCAATGCTGATGAGTACAGAGTAAATGATGCGGCTGCG 736
|||
Db 361 GCGTGGCTGGGGCTTCAATGCTGATGAGTACAGAGTAAATGATGCGGCTGCG 420
|||
Qy 737 GGGAGATGCGCTTACCTGCTGAGACCTTCTGCTGCGAGCGCTGCGACCTACCGCG 796
|||
Db 421 GGGAGATGCGCTTACCTGCTGAGACCTTCTGCTGCGAGCGCTGCGACCTACCGCG 480
|||
Qy 797 GCGCACTGCTGCTGCTGCTGAGAGAGTGGGTTCAAGACACTGCTGCTGCTGCTG 856
|||
Db 481 GCGCACTGCTGCTGCTGCTGAGAGAGTGGGTTCAAGACACTGCTGCTGCTGCTG 540
|||
Qy 857 TGGTCCACAGAGGGCGGCTAGATGACACCTGCTGCTGCTGAGAGAGAGTGGCTTCTC 916
|||
Db 541 TGGTCCACAGAGGGCGGCTAGATGACACCTGCTGCTGCTGAGAGAGAGTGGCTTCTC 600
|||
Qy 917 ACTGCAAGTGGTGGCGGAGAGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
|||
Db 601 ACTGCAAGTGGTGGCGGAGAGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
|||
Qy 977 TGG--CGAGTGAATGTTTGGTGCACCCATGAGG--CGCTGTGACAGTGGCGTCAATG 1033
|||
Db 661 TGGGACAGTGAATGTTTGGTGCACCCATGAGG--CGCTGTGACAGTGGCGTCAATG 720
|||
Qy 1034 TGTCTACCTGCGGGCGG-----GAGTGTCTCGTGCATGCTGCTGCGGCTCTTCCCTG 1087
|||
Db 721 TGTCTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 780
|||
Qy 1088 CCAAT--CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1139
|||
Db 781 NCCTTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
|||
Qy 1140 GCGCGTGGTGGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1195
|||
Db 841 GCGCGTGGGCGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
|||
Qy 1196 ATGGGCAATGATGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1239
|||
Db 901 AATGGGGGAAATGAATGTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 944
|||

```

RESULT 9
AM963415/c 676 bp mRNA linear EST 01-JUN-2000
LOCUS EST3737488 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AM963415
VERSION AM963415.1 GI:8153251
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 676)
Hegde, P., Ol, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt
I., E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
Unpublished (2000)
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 187

```

FEATURES                               Seq primer: Reverse.
Source                                  Location/Qualifiers
1..676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="MAGE resequences, MAGE"
/note="Vector: pBluescriptSM"

BASE COUNT      147 a      220 c      188 g      121 t

Query Match      40.1%; Score 641.4; DB 10; Length 676;
Best Local Similarity 98.2%; Pred. No. 7.1e-131;
Matches 659; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

Qy 927 GGTCCCGGAGAGAGACTTGTGCTAG--TCTGACACTGCTGACACTGACACTGACACTG 985
|||
Db 671 GGGGCCCCAGAGAGACTTGTGCTAGTTCGGGCAATTGGGCACTTGCACCTGGGACAGTG 612
|||
Qy 986 AGATGTTGCTGACACCATGGGCGGCTGCTGACACTGCGCTGACATGCTGCTGCTGCTG 1045
|||
Db 611 AGATGTTGCTGACACCATGGGCGGCTGCTGACACTGCGCTGACATGCTGCTGCTGCTG 552
|||
Qy 1046 GGGGCGGAGTGTCTGCGGCGCATGCTGCTGCGGCTTCTGCTGACATGAGAGAGAGAG 1105
|||
Db 551 GGGGCGGAGTGTCTGCGGCGCATGCTGCTGCGGCTTCTGCTGACATGAGAGAGAGAG 492
|||
Qy 1106 ATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1165
|||
Db 491 ATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
|||
Qy 1166 CCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1225
|||
Db 431 CCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372
|||
Qy 1226 AGAGCGAGTGCACCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGAT 1285
|||
Db 371 AGAGCGAGTGCACCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGAT 312
|||
Qy 1286 GCTGGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1345
|||
Db 311 GCTGGAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252
|||
Qy 1346 TGGCTAGTGTCTACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1405
|||
Db 251 TGGCTAGTGTCTACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 192
|||
Qy 1406 AGCTCTGTGGGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1465
|||
Db 191 AGCTCTGTGGGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132
|||
Qy 1466 TTGGGGGAGACAGGCAAGATGAAGTCTTGGGTGAGAGAGAGAGAGAGAGAGAGAGAG 1525
|||
Db 131 TTGGGGGAGACAGGCAAGATGAAGTCTTGGGTGAGAGAGAGAGAGAGAGAGAGAGAG 72
|||
Qy 1526 CCTATGTAGAGGCTTCTGATGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1585
|||
Db 71 CCTATGTAGAGGCTTCTGATGTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12
|||
Qy 1586 AGTGACATTCC 1596
|||
Db 11 AGTGACATTCC 1


```

RESULT 10
BE464487/c 702 bp mRNA linear EST 27-JUL-2000
LOCUS h21710.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:3208290 3'
DEFINITION Similar to TR:088886 088886 SPHINGOSINE KINASE. ;, mRNA sequence.
ACCESSION BE464487
VERSION BE464487.1 GI:9510262
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

through R-DC0-CBA, R-DD0-B2R through R-DD0-CAA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. 9) a pool of about 2,000 clones from the pool of normalized libraries, CN0, that makes up the tracer. The corresponding plates are R-CN0-BKW through R-CN0-BLD, R-CN0-BLG, R-CN0-BLP through R-CN0-BLR, R-CN0-BLT, R-CN0-BLM-X, R-CN0-BMB, and R-CN0-BMF through R-CN0-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkw-a-09-0-UI, bkw-b-09-0-UI, bkw-b-11-0-UI, bkw-b-10-0-UI, bkw-d-01-0-UI, bkw-d-06-0-UI, bkw-g-08-0-UI, bkw-h-12-0-UI, bkw-a-05-0-UI, bkw-a-06-0-UI, bkw-a-11-0-UI, bkw-c-06-0-UI, bkw-c-09-0-UI, bkw-d-10-0-UI, bkw-a-01-0-UI, bkw-a-02-0-UI, bkw-f-04-0-UI, bkw-g-07-0-UI, bkw-g-12-0-UI, bkw-h-12-0-UI, bkw-f-02-0-UI, bkw-f-11-0-UI, bkw-f-04-0-UI, bkw-e-95-0-UI, bkw-d-08-0-UI, bkw-f-02-0-UI, bkw-h-04-0-UI, bkw-a-05-0-UI, bkw-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CN0 clone (corresponding to the address bkw-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG_SEQ=None found"

BASE COUNT 136 a 210 c 178 g 138 t

Query Match 38.5%; Score 615.8; DB 13; Length 662;
Best Local Similarity 98.7%; Pred. No. 3e-125;
Matches 631; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```

Qy 961 CTGTGCACTGGACCTGGCCAGATGATGTTTCTGCAACCCAGCCGCTGTGCACT 1020
Db 654 CGGACAGAGGGCCACTGGGACATGATG-TTCTGCAACCCAGCCGCTGTGCACT 596
Qy 1021 GGCCTCATGATCTGTTCTACTAGTCGGGGGAGTGTCTGTCGATGCTGCGGCTC 1080
Db 595 GGCCTCATGATCTGTTCTACTAGTCGGGGGAGTGTCTGTCGATGCTGCGGCTC 536
Qy 1081 TTCTGGCCATGGAAGAGGAGCATATGAGATGATGATGATGATGATGATGATG 1140
Db 535 TTCTGGCCATGGAAGAGGAGCATATGAGATGATGATGATGATGATGATGATG 476
Qy 1141 CCCGTGGTGGCTTCCGCTTGGAGCCCAAGATGGAAGAGTGTGTTGCAAGTGAAG 1200
Db 475 CCCGTGGTGGCTTCCGCTTGGAGCCCAAGATGGAAGAGTGTGTTGCAAGTGAAG 416
Qy 1201 GAATTGATGTTTACGAGGCGGTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Db 415 GAATTGATGTTTACGAGGCGGTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 356
Qy 1261 AGCGGTGCGTGAAGCCCGCCAGCTGGAAGCCCAAGAGGAGGAGGAGGAGGAGG 1320
Db 355 AGCGGTGCGTGAAGCCCGCCAGCTGGAAGCCCAAGAGGAGGAGGAGGAGGAGG 296
Qy 1321 CCCTTATGACCCCTGGGCGCGGCTGTGCTTATGCTTACTTGAAGACCCCTTCT 1380
Db 295 CCCTTATGACCCCTGGGCGCGGCTGTGCTTATGCTTACTTGAAGACCCCTTCT 236
Qy 1381 CCCTAGGAGCTGAGGCGCTGTCACAGCTCTGTGGGGGAGGAGGAGGAGGAGGAG 1440
Db 235 CCCTAGGAGCTGAGGCGCTGTCACAGCTCTGTGGGGGAGGAGGAGGAGGAGGAG 176
Qy 1441 AGGCTGAGAAAGTGAAGCTATGCTTGGGGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Db 175 AGGCTGAGAAAGTGAAGCTATGCTTGGGGGAGGAGGAGGAGGAGGAGGAGGAG 116
Qy 1501 GAGCCCAAGCTGGGCGGAGCTGCTATGTAAGGCTTCTGTTGTTGTTGAGACCC 1560
Db 115 GAGCCCAAGCTGGGCGGAGCTGCTATGTAAGGCTTCTGTTGTTGTTGAGACCC 56

```

Qy 1561 CACCCAGAGAACCAATCAATTAAGTGAATTCCTCCAA 1599
Db 55 CACCCAGAGAACCAATCAATTAAGTGAATTCCTCCAA 17

RESULT 13
LOCUS BM559257
DEFINITION BM559257 1005 bp. mRNA linear EST 20-FEB-2002
ACCESSION AGNCNCURT_6598353 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5474413
VERSION 5, mRNA sequence.
KEYWORDS BM559257.1 GI:18802694
SOURCE EST.

ORGANISM

human.

EuKaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1005)

AUTHORS

NIH-MGC <http://mhc.ncl.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF

FEATURES

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: LNCM1987 row: j column: 14
High quality sequence stop: 618.

FEATURES

Location/Qualifiers

1..1005
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5474413"
/clone_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant), cell line"

/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 157 a 333 c 348 g 167 t

Query Match 38.4%; Score 614.2; DB 13; Length 1005;
Best Local Similarity 94.4%; Pred. No. 7.3e-125;
Matches 671; Conservative 0; Mismatches 33; Indels 7; Gaps 3;

```

Qy 3 CCGCGGGCTCTTATGACGAGCTCCGGGGGAGGAGGAGGAGGAGGAGGAGGAGG 62
Db 154 CCGCGGGCTCTTATGACGAGCTCCGGGGGAGGAGGAGGAGGAGGAGGAGGAGG 213
Qy 63 CGAGCCCGCTTGGAGCAGCAGTAAAGAGCTGAAGCAGAGACCCGCGCAGGCGAG 122
Db 214 CGAGCCCGCTTGGAGCAGCAGTAAAGAGCTGAAGCAGAGACCCGCGCAGGCGAG 273
Qy 123 CGCCCCACAGCGCCAGGAGCCCTTGCGACCGGAGCCGCGGCTGAGATTATGATCC 182
Db 274 CGCCCCACAGCGCCAGGAGCCCTTGCGACCGGAGCCGCGGCTGAGATTATGATCC 333
Qy 183 AGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 242
Db 334 AGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 393
Qy 243 GCGGGGGGGAAGGAGGAGGCTTTCAGAGCTTTCGAGTACGTGAGCCCTTTTGGC 302

```


Db		394	GCCGGCGGCAAGGCGCCTTCACGCTCTTCCGAGTCAAGTGACAGCCCTTTTGGC	453
Oy		303	TGAAGGTGAATCTCTTTCACGCTGATGCTCACTGAGCGGGACAACCAGCGCGGAGCT	362
Db		454	TGAGGCTGAATCTCTTTCACGCTGATGCTCACTGAGCGGGACAACCAGCGCGGAGCT	513
Oy		363	GGTGCAGTGGAGAGACTGGCGCGCTGGAGCGCTGGTGTATGCTCTGAGAGCGAGCT	422
Db		514	GGTCCGGTGGAGAGAGTGGCGCGCTGGAGCGCTGGTGTATGCTCTGAGAGCGAGCT	573
Oy		423	GATGCACAGAGTGGTGAACGGGGCTCATGAGACGGCGCTGACTGGGAGACGCCCATCCAGA	482
Db		574	GATGCACAGAGTGGTGAACGGGGCTCATGAGAGCGGCTGACTGGGAGACGCCCATCCAGA	633
Oy		483	GCCCCGTGTAGCTCCCGACAGAGCTCTGGCAACGGCGCTGGCACCTCTTGAACCATTA	542
Db		634	GCCCCGTGTAGCTCCCGACAGAGCTCTGGCAACGGCGCTGGCACCTCTTGAACCATTA	693
Oy		543	TGCTGGCATAGAGAGGTCAACCAATGAAGACCTCTGACCAACTGCAACGCTATTGCTGTG	602
Db		604	TGCTGGCATAGAGAGGTCAACCAATGAAGACCTCTGACCAACTGCAACGCTATTGCTGTG	753
Oy		603	CGGCCCG-GTGCCTGCACCATG-AACCTGCTGTCTGCGACAGAGGCTTGGGGGCTGGC	660
Db		754	CGCGCGGCTGTGTCACCATGAAACCTGTGCTGTGACACAGGGGCTTGGGGGCTGGC	813
Oy		661	TGCTTCTCMGTGCTCAGGC-----TGCGCTGGGGCTTATGCTGANTGTG	706
Db		814	CCTCCTTCTCTGCTCCTCCACCCCTGGCCTGGGGCTTATGCTGAATGGG	864
RESULT 14				
BG280830		809 bp	mRNA	linear EST 21-FEB-2001
LOCUS		602401209P1 NIH_MGC_20 Homo sapiens CDNA clone IMAGE:4543395 5'		
DEFINITION		mRNA sequence.		
ACCESSION		BG280830		
VERSION		BG280830.1 GI:13029767		
KEYWORDS		EST.		
SOURCE		human.		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE		1 (bases 1 to 809)		
JOURNAL		NIH-MGC http://mgc.ncl.nih.gov/		
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
		Contact: Robert Strausberg, Ph.D.		
		Email: cgaabs-r@mail.nih.gov		
		Tissue Procurement: ATCC/DICP		
		cDNA Library Preparation: Ling Hong/Rubin Laboratory		
		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
		DNA Sequencing by: Incyte Genomics, Inc.		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov		
		Plate: ILICMH225 row: b column: 04		
		High quality sequence stop: 688.		
FEATURES				
source		1..809		
		Location/Qualifiers		

1. .809
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4543395"
/clone_1b="NH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOMB7; Site:1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5', insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT	165 a	240 c	250 g	153 t	1 others
ORIGIN					
Query Match	38.0%;	Score 608.4;	DB 12;	Length 809;	
Best Local Similarity	95.8%;	Pred. No. 1,3e-123;			
Matches 710;	Conservative	0;	Mismatches 22;	Indels 9;	Gaps 8;
QY	868	GGCCCCGTAAATGACACACCGTGGTCCACATGAGAGACAGAGTGCCTTCTACATGAGCAGGTG	927		
DB	7	GGCCCCGTAAATGACACACCTTGGTCCACATGAGAGACAGAGTGCCTTCTACATGAGCAGGTG	66		
QY	928	GTGCCCCAGCAGAGACTTGTGTAGTCTGGCAGCTCTGCACTCTGCACTCTGCACTCTGCACTG	987		
DB	67	GTGCCCCAGCAGAGACTTGTGTAGTCTGGCAGCTCTGCACTCTGCACTCTGCACTCTGCACTG	126		
QY	988	ATGTTTGTCTGACACCAATGAGGCGCGTGTGACAGAGGAGGAGTGCATGTGTTCTACAGTGGG	1047		
DB	127	ATGTTTGTCTGACACCAATGAGGCGCGTGTGACAGAGGAGTGCATGTGTTCTACAGTGGG	186		
QY	1048	GCAGGAGTGTCTGTCGCATGCTGCTGCGCTCTCTTCTGCGCATGAGAGAGGAGGAGCAT	1107		
DB	187	GCAGGAGTGTCTGTCGCATGCTGCTGCGCTCTCTTCTGCGCATGAGAGAGGAGGAGCAT	246		
QY	1108	ATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1167		
DB	247	ATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	306		
QY	1168	AAGGATGAGGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1227		
DB	307	AAGGATGAGGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	366		
QY	1228	GGCCAGGTGACACCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT	1287		
DB	367	GGCCAGGTGACACCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT	426		
QY	1288	TGGAAGCCCCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1347		
DB	427	TGGAAGCCCCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	486		
QY	1348	CCTTAGTGTCTACTTTCAGAGAGACCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1407		
DB	487	CCTTAGTGTCTACTTTCAGAGAGACCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	546		
QY	1408	CCTTAGT	1463		
DB	547	CCTTAGT	606		
QY	1464	CTTTTG-GGGGGGACAGAGCCA-GAATGAAGTCTGTGGGTACAGAGAGCCAGTGGCTGGGCCCA	1521		
DB	607	CTTTTG-GGGGGGACAGAGCCA-GAATGAAGTCTGTGGGTACAGAGAGCCAGTGGCTGGGCCCA	666		
QY	1522	-GCTGCCATATGTAAGGCTCTTCTAG-TTGTGTGTGAGAC-CCACCCACAGACCAACATC	1578		
DB	667	GCTGCCATATGTAAGGCTCTTCTAG-TTGTGTGTGAGAC-CCACCCACAGACCAACATC	726		
QY	1579	CAATTAAGTGAATTCACCA 1599			
DB	727	CAATTAAGTGAATTCACCA 747			
RESULT 15					
LOCUS	BM557357	1075 bp	mRNA	linear	EST 20-FEB-2002
DEFINITION	ABENCOUR.6578959 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466709				
ACCESSION	BM557357				
VERSION	BM557357.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1075)				

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LICM1967 row: 1 column: 14
High quality sequence start: 39
High quality sequence stop: 680.
Location/Qualifiers

FEATURES
SOURCE
1. 1075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5466709"
/clone_lid="NIH-MGC.41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 174 a 354 c 382 g 164 t 1 others
ORIGIN

Query Match 37.9%; Score 606.8; DB 13; Length 1075;
Best Local Similarity 97.7%; Pred. No. 3.1e-123;
Matches 647; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 3 CCCCCGGGCTCTTATACCGAGGCTCCGGGGGGGGAAGCGCAGCCCAACAGCCGGCCCTG 62
DB 211 CGCCGGTCTCTACACCGAGGCTCCGGGGGGGGAAGCGCAGCCCAACAGCCGGCCCTG 270
QY 63 CGAGCGCCGCTGGGAGACCGATGAAGAGCTGAAGAGAGCGCCGCGCACGGGCGAG 122
DB 271 CGAGCGCCGCTGGGAGACCGATGAAGAGCTGAAGAGAGCGCCGCGCACGGGCGAG 330
QY 123 CGCCCGCACAGCGCCAGGAGCCCTGGAGCGGAGCCGCGGTGAGATTATGATCC 182
DB 331 CGCCCGCACAGCGCCAGGAGCCCTGGAGCGGAGCCGCGGTGAGATTATGATCC 390
QY 183 AGCGGG 242
DB 391 AGCGGG 450
QY 243 GCGGGGGGCAAGGGCAAGGGCTTGACAGCTTCCGGAGTACGTGACGAGCCCTTTGGG 302
DB 451 GCGGGGGGCAAGGGCAAGGGCTTGACAGCTTCCGGAGTACGTGACGAGCCCTTTGGG 510
QY 303 TGAGGCTGAATCTCTTACAGCTGATGCTCAGTGAAGCGGGA - ACCACGCGGGAGC 361
DB 511 TGAGGCTGAATCTCTTACAGCTGATGCTCAGTGAAGCGGGA - ACCACGCGGGAGC 570
QY 362 TGGTGGGCTGGAGAGCTGGGGCGGCTGGGAGCGCTGTGGTATGTCTGGAGAGCGGC 421
DB 571 TGGTGGGCTGGAGAGCTGGGGCGGCTGGGAGCGCTGTGGTATGTCTGGAGAGCGGC 630
QY 422 TGATGACGAGAGTGTGAACGGGCTCATGAGCGGCTGACTGGAGACCGCATCCAGA 481
DB 631 TGATGACGAGAGTGTGAACGGGCTCATGAGCGGCTGACTGGAGACCGCATCCAGA 690
QY 482 AGCCCTGTGTAGCTTCCACAGAGCTGTGGCAAAGCGCTGGACAGCTTCTTGAACATT 541
DB 691 AGCCCTGTGTAGCTTCCACAGAGCTGTGGCAAAGCGCTGGACAGCTTCTTGAACATT 750

QY 542 ATGCTGCTATGACACGAGTCAACCAATGAGAAGCTCTGACCAACTGACGCTATTGCTGT 601
DB 751 ATGCTGCTATGACACGAGTCAACCAATGAGAAGCTCTGACCAACTGACGCTATTGCTGT 810
QY 602 GCCGCCCGG-TGCTGTACCCAT-GAACCTGCTGTCTGTGCACAGCGCTTCGGGGCTGCG 659
DB 811 GCCGCCCGGCTGTGTACCCGTTGAACCTGTCTGTGTGCACAGCGGCTTCGGGGCTGCG 870
QY 660 CT 661
DB 871 CT 872

Search completed: July 12, 2003, 15:59:46
Job time : 2390 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 13:30:22 ; Search time 82 Seconds
(Without alignments)
5980.197 Million cell updates/sec

Title: US-09-784-810A-1_COPY_2_1600
Perfect score: 1599
Sequence: 1 cccccggggctctctatagc.....aataaagtacattcccaa 1599

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5b_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCtUS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.6	3.1	1203	US-09-086-010-1	Sequence 1, Appl1
2	48.6	3.0	4403765	US-09-103-840A-2	Sequence 2, Appl1
3	47.8	3.0	900	PCT-US95-04801-3	Sequence 3, Appl1
4	47.2	3.0	2793	US-08-209-747-1	Sequence 1, Appl1
5	47.2	3.0	2793	US-08-458-288-1	Sequence 1, Appl1
6	47.2	2.9	68750	US-09-335-409-1	Sequence 1, Appl1
7	47.2	2.9	68750	US-09-568-102-1	Sequence 1, Appl1
8	47.2	2.9	68750	US-09-567-969-1	Sequence 1, Appl1
9	47.2	2.9	68750	US-09-568-480-1	Sequence 1, Appl1
10	47.2	2.9	68750	US-09-568-486-1	Sequence 1, Appl1
11	47.2	2.9	68750	US-09-568-472-1	Sequence 1, Appl1
12	47.2	2.9	68750	US-09-567-899-1	Sequence 1, Appl1
13	46.6	2.9	2011	US-09-567-899-1	Sequence 1, Appl1
14	46.4	2.9	53526	US-08-658-136-2	Sequence 2, Appl1
15	46.4	2.9	53577	US-08-658-136-1	Sequence 2, Appl1
16	45.5	2.8	4411529	US-09-103-840A-1	Sequence 1, Appl1
17	44.8	2.8	1164	US-07-640-476-6	Sequence 6, Appl1
18	44.6	2.8	4257	US-08-690-473-1	Sequence 1, Appl1
19	44.6	2.8	4257	US-09-259-821A-1	Sequence 1, Appl1
20	44.6	2.8	4257	US-08-843-659-1	Sequence 1, Appl1
21	44.6	2.8	12001	US-08-458-568A-11	Sequence 11, Appl1
22	44.2	2.8	4403765	US-09-103-840A-2	Sequence 2, Appl1
23	43.8	2.7	71989	US-09-443-501A-2	Sequence 2, Appl1
24	43.4	2.7	1921	US-08-482-677-3	Sequence 3, Appl1
25	43.4	2.7	1926	US-08-152-019A-43	Sequence 43, Appl1
26	43.3	2.7	28804	US-08-592-874-1	Sequence 1, Appl1
27	43.3	2.7	28804	US-09-096-942-2	Sequence 2, Appl1

28	43	2.7	28804	3	US-09-096-867-2	Sequence 2, Appl1
29	42.8	2.7	43280	2	US-08-804-227C-1	Sequence 1, Appl1
30	42.8	2.7	44377	2	US-08-804-227C-7	Sequence 1, Appl1
31	42.8	2.7	44377	2	US-08-804-198-1	Sequence 1, Appl1
32	42.8	2.7	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
33	42.6	2.7	2448	4	US-08-487-596-13	Sequence 9, Appl1
34	42.6	2.7	2450	2	US-08-466-589-9	Sequence 9, Appl1
35	42.6	2.7	2450	2	US-08-700-636-9	Sequence 9, Appl1
36	42.6	2.7	2450	3	US-08-467-574-9	Sequence 9, Appl1
37	42.6	2.7	2450	3	US-09-217-345-9	Sequence 9, Appl1
38	42.4	2.7	530	3	US-08-758-662-4	Sequence 4, Appl1
39	42.4	2.7	4524	2	US-08-845-998-7	Sequence 7, Appl1
40	42.4	2.7	4524	3	US-09-206-537-7	Sequence 7, Appl1
41	42.4	2.7	4524	4	US-09-430-854-7	Sequence 7, Appl1
42	42.2	2.6	7271	6	5215881-2	Sequence 7, Appl1
43	42.2	2.6	8438	1	US-07-945-283-1	Patent No. 5215881
44	42.2	2.6	1131	4	US-09-420-211-1	Sequence 1, Appl1
45	42	2.6	1843	1	US-07-918-023-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-086-010-1
Sequence 1, Application US/09086010
Patent No. 6274338
GENERAL INFORMATION:
APPLICANT: Glmcher, Laurie H. et al.
TITLE OF INVENTION: Human c-Maf Compositions and
TITLE OF INVENTION: Methods of use thereof
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/086,010
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/030,579
FILING DATE: 2-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: 41,106
REFERENCE/DOCKET NUMBER: HUI-027CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1203 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1203
US-09-086-010-1
Query Match 3.1%, Score 49.6; DB 4; Length 1203;
Best Local Similarity 45.6%; Pred. No. 0.016;
Matches 175; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 2 CCCCCGGGGCTCTCTATAGCCAGGCTCCGGGGGGGGAAGCGACGCCCAAGAGCCGCCCT 61
DB 494 CCGCCCTGATTCGCGCGCGCGCGCGCAAGCGGGGGCGCCGCACTACACACACAC 553
QY 62 GCGAGCCCGCCCTGGGCGAGCAGATAGAGCTGAAAGCAGAGAGCCGCCACAGGCA 121
DB 554 ACCACACAGCCCGCGCGCGCGCGCAACACCGAGCGCGCGCGCGCGCGCGCGCA 613
QY 122 GCGCCCGCCACAGCGAGGAGCCCTTGCAGCGGGAGCGCGGGGTCAGATTAGATC 181
DB 614 GCGCGCGCGCTTGGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 673
QY 182 CAGCGGGGGCG 241
DB 674 GGGGGGGGGCG 733
QY 242 CCGCGCGCGCGCGAGGCGAGGCGCTTGCAGCTTTCGGAGTCACTGCGAGCCCTTTGG 301
DB 734 CGCACACAGCG 793
QY 302 CTGAGGCGTGAATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
DB 794 CCAATGCTGCTGCGGAGTGAACCGGAGCTGGCGGGGGGTACAGAGGAGAGTATCC 853
QY 362 TGGTGGCGTGGAGGAGCTGGGCC 385
DB 854 GGCTGAAGCAGAGAGCGCGACCC 877

RESULT 2

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match

Best Local Similarity 47.2%; Score 48.6; DB 4; Length 4403765;
Matches 119; Conservative 0; Mismatches 199; Indels 1; Gaps 1;

QY 2 CCCCCGGGGCTCTCTATAGCCAGGCTCCGGGGGGGGAAGCGACGCCCAAGAGCCGCCCT 61
DB 3942920 CCGCCGAGGCT 3942861
QY 62 GCGAGCGCGCGCTGGGAGCAGATAGAGCTGAAGCAGAGCGCGCGCGCGCGCGCA 121
DB 3942860 CCGCGCTTACACCGTTACACCGTTACACCGTTACACCGTTACACCGCGG 3942801
QY 122 GCGCGC-CCACAGCGCGAGGAGCCCTTGGCAGCGGAGCGCGGGTCAAGATTAGAT 180
DB 3942800 GCGCGCTCGTGGGCGCATGCGCGCGCTACCGCGCGCGCGCGCTTGGCGCG 3942741
QY 181 CCAAGCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTG 240

DB 3942740 CTGCGGTGGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCTGCGCG 3942681
QY 241 CCGCGCGCGCGCAAGGCGCAAGGCGCTTGCAGCTTTCGGAGTACGTGACGCGCTTTT 300
DB 3942680 CCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942621
QY 301 GCTAGAGCTGAATCTCTTCACTGAGTGTCTACTGAGCGGCGGAACACGCGCGGAG 360
DB 3942620 CTGCG 3942561
QY 361 CTGTCGCGTGGAGAGAGC 379
DB 3942560 CCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942542

RESULT 3

PCT-US95-04801-3
; Sequence 3, Application PC/TUS9504801
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan F.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enguita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Liarena, Francisco J.
; APPLICANT: Liras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHARYCIN BIOSYNTHESIS
; TITLE OF INVENTION: LATE GENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04801
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US95-04801-3

Query Match

Best Local Similarity 51.1%; Score 47.8; DB 5; Length 900;
Matches 112; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 184 GCGGCG 243
DB 353 GCGGCG 412
QY 244 GCGGCG 303
DB 413 GCGAGCG 472

QY 304 GAGCGTGAATCTCTTACGCTGATGCTACCTAGAGCGGGAACCAACGCGGAGCTG 363
DB 473 CAGATGATCAACGCTGGCGCGGCTGCTGAGACCGGCTGGCAACTTCCGCTAGGCGCTG 532
QY 364 GTGCGGTCCGAGAGACTGGGCGCTGGAGACCTCTGTG 402
DB 533 TGGCACTACCGGTCTGCTGACCTGTCCGCGCACTGTGTG 571

RESULT 4

US-08-209-747-1
Sequence 1, Application US/08209747
Patent No. 5733771
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: CDNAs Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,747
FILING DATE: 14-MAR-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
TISSUE TYPE: minor ampullate gland
FEATURE:
NAME/KEY: CDS
LOCATION: 183..2675
OTHER INFORMATION: /product= "N. clavipes minor
OTHER INFORMATION: ampullate silk protein"
US-08-209-747-1

Query Match 3.0%; Score 47.2; DB 1; Length 2793;
Best Local Similarity 47.1%; Pred. No. 0.067;
Matches 145; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 150 GCACGGGAGCCGCGGCTCAGATTATGATCCAGCGGCGGCGCGGCGGCTGCTCC 209
DB 1331 GCAGAGACCCCTGCGGGCTGACGAGAGCTACGCTACGCTCAAGGTGGTACGGTCC 1390
QY 210 GCGGCGCTGCGGGTGTGCTGCTGCTGAACCGCGGCGGCGGCAAGGCTTGA 269
DB 1391 GGTGACAGAGCTGTGCTGGCTGCTGCTGAGACAGAGCTGAGGCGCTGCTGTAC 1450
QY 270 GCTCTTCGAGATCAGCTGACGCCCTTTGGCTGAGGCTGAATCTCTTACAGCTGAT 329

DB 1451 GGTAGAGGTCTGTGCTGAGCTGAGCTGACAGCGCCAGAGCTGAGAGCTACGGT 1510
QY 330 GCTCAGTACGAGCGGAGAACACCGCGGAGAGCTGCTGGTCCGAGAGAGCTGGGCGCTG 389
DB 1511 GGTCAAGGTGATACGCTGCGGAGCAGAGCTGCTGCTGAGAGCTGCTGCAACAGGA 1570
QY 390 GGAGCTCTGTGATCTGCTGAGACGAGCTGATGACAGAGTGTGAACGGCTCAT 449
DB 1571 GCCGAGCGCTGTGTGCTTACGGTAGAGGTGTGCTGTGAGCTGTGCGCGCTGCTGG 1630
QY 450 GGAGCGGC 457
DB 1631 GCAGTGC 1638

RESULT 5

US-08-458-298-1
Sequence 1, Application US/08458298
Patent No. 5756677
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Colgin, Mark
TITLE OF INVENTION: CDNAs Encoding Minor Ampullate Spider
TITLE OF INVENTION: Silk Proteins
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,298
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,747
FILING DATE: 14-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2793 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
TISSUE TYPE: minor ampullate gland
FEATURE:
NAME/KEY: CDS
LOCATION: 183..2675
OTHER INFORMATION: /product= "N. clavipes minor
OTHER INFORMATION: ampullate silk protein"
US-08-458-298-1

Query Match 3.0%; Score 47.2; DB 1; Length 2793;
Best Local Similarity 47.1%; Pred. No. 0.067;
Matches 145; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

PA (GETH) GENENTECH INC.
 XX Rastelli L;
 PI WPI: 2001-514770/56.
 DR P-PSDB; AAE07882.
 XX An isolated sphingosine kinase polypeptide useful for treating a
 PT SphK-associated disorder especially cancer, restenosis or ischemia in a
 PT human -
 PS Claim 8; Page 89-90; 107pp; English.
 XX The present invention relates to sphingosine kinase (SphK) polypeptides
 CC and nucleic acids encoding them. SphK is useful for treating a sphk-
 CC associated disorder especially cancers such as leukaemia, lymphoma,
 CC ovarian, breast, lung, colon, testicular, stomach and skin,
 CC atherosclerosis, restenosis or ischemia and cell proliferative disease
 CC or disorder associated with vascular diseases. SphK gene is used in gene
 CC therapy and antisense therapy. Sphingolipids serving as signalling
 CC molecules, have recently emerged as regulators of cell growth,
 CC differentiation, diverse cell phenotypes and cell death. Activation of
 CC SphK by tumour necrosis factor (TNF)-alpha inhibits apoptosis in human
 CC endothelial cells. The present sequence is human sphingosine kinase
 CC (SphK) cDNA.
 SQ Sequence 1600 BP; 265 A; 492 C; 531 G; 311 T; 1 other:
 Query Match 100.0%; Score 1599; DB 22; Length 1600;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCCCGGGGCTCTATAGCCAGGCTCCGGGGGAGGCGGCCACAGCGGGCCC 60
 2 CCCCCGGGGCTCTATAGCCAGGCTCCGGGGGAGGCGGCCACAGCGGGCCC 61
 61 TCGAGCGCCGCTGGGCGACGATAGAGAGTGAAGAGAGAGCGGCCACAGGCG 120
 62 TCGAGCGCCGCTGGGCGACGATAGAGAGTGAAGAGAGAGCGGCCACAGGCG 121
 121 AGGCGCCCAAGAGCGCCAGGAGCCCTGGAGCGGAGAGCGCGGGGTGAGTATGAT 180
 122 AGGCGCCCAAGAGCGCCAGGAGCCCTGGAGCGGAGAGCGCGGGGTGAGTATGAT 181
 181 CCAAGCGGGCGCCCGGGGGCTGTCCGGGCGCTGCGGCTGTGTGTGTAAC 241
 182 CCAAGCGGGCGCCCGGGGGCTGTCCGGGCGCTGCGGCTGTGTGTGTAAC 241
 241 CCGGGGGGGGCAAGGAGCGCTTGCAGCTTCTCCGAGTACAGTGCAGCCCTTTTG 300
 242 CCGGGGGGGGCAAGGAGCGCTTGCAGCTTCTCCGAGTACAGTGCAGCCCTTTTG 301
 301 GCTGAGCTGAATCTCTTCAAGCTGATGCTACAGTGAAGGAGGAGCAAGCGGGAG 360
 302 GCTGAGCTGAATCTCTTCAAGCTGATGCTACAGTGAAGGAGGAGCAAGCGGGAG 361
 361 CTGGGCGGTGGAGAGTGGGCGGCTGGAGCGTGTGTGATGTGTGTAACGGG 420
 362 CTGGGCGGTGGAGAGTGGGCGGCTGGAGCGTGTGTGATGTGTGTAACGGG 421
 421 CTGATGACGAGAGTGTGTAAGGAGCTGATGAGGAGGCTGATGAGGAGCGGCAATCAG 480
 422 CTGATGACGAGAGTGTGTAAGGAGCTGATGAGGAGGCTGATGAGGAGCGGCAATCAG 481
 481 AAGCCCTGTGTAGCTCCCAAGAGCTTGGCAAGCGGTGGCAAGCTTCTTGAACAT 540
 482 AAGCCCTGTGTAGCTCCCAAGAGCTTGGCAAGCGGTGGCAAGCTTCTTGAACAT 541
 541 TATGCTGCTATGACAGGATCAACATTAAGACTCTGACCACTGACAGCTATGCTG 600
 542 TATGCTGCTATGACAGGATCAACATTAAGACTCTGACCACTGACAGCTATGCTG 601
 601 TGCCTCCCGGCTGTCAACCCATGAACCTGTCTGTCAACAGGCTTGGGGCTGCGC 660

602 TGCCTCCCGGCTGTCTACACCATGAACCTGTCTGTCAACAGGCTTGGGGCTGCGC 661
 661 TGCCTCCCGGCTGTCTACACCATGAACCTGTCTGTCAACAGGCTTGGGGCTGCGC 720
 662 TGCCTCCCGGCTGTCTACACCATGAACCTGTCTGTCAACAGGCTTGGGGCTGCGC 721
 721 AAGTATCGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCTGCTGGAGCC 780
 722 AAGTATCGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCTGCTGGAGCC 781
 781 CTGGGAGCTTACCGCGGAGCTGCTACCTCCCTGTAGAGAGTGGGTTTCAAGCA 840
 782 CTGGGAGCTTACCGCGGAGCTGCTACCTCCCTGTAGAGAGTGGGTTTCAAGCA 841
 841 CCGTCTCCCGGCTGTGTGTCAAGAGGCGCGGTGTAGAGAGTGGGTTTCAAGCA 900
 842 CCGTCTCCCGGCTGTGTGTCAAGAGGCGCGGTGTAGAGAGTGGGTTTCAAGCA 901
 901 GAGCAGGAGCTTCTCACTGAGAGTGTGTCGAGAGAGTGTGTGTGTGTGTGTGTGT 960
 902 GAGCAGGAGCTTCTCACTGAGAGTGTGTCGAGAGAGTGTGTGTGTGTGTGTGTGT 961
 961 CTGCTGACCTGACCTGCGCAGTGAAGATTTGTGACACCATGGGCGCTGTGACGT 1020
 962 CTGCTGACCTGACCTGCGCAGTGAAGATTTGTGACACCATGGGCGCTGTGACGT 1021
 1021 GAGCTGACATCTGTCTTACGTGCGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 1022 GAGCTGACATCTGTCTTACGTGCGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1081
 1081 TTCTGCGCATGAGAGAGGAGGAGATGAGATGATGATGATGATGATGATGATGATG 1140
 1082 TTCTGCGCATGAGAGAGGAGGAGATGAGATGATGATGATGATGATGATGATGATG 1141
 1141 CCGTGTGCGCTTCCGCTTGGAGCCAGAGATGGAAGATGGAAGATGGAAGATGGAAG 1200
 1142 CCGTGTGCGCTTCCGCTTGGAGCCAGAGATGGAAGATGGAAGATGGAAGATGGAAG 1201
 1201 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 1202 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1261
 1261 AGCGGT 1320
 1262 AGCGGT 1321
 1321 CCCTTATGACCTCTGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
 1322 CCCTTATGACCTCTGGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1381
 1381 CCGTAGGCTGTGAGGAGGCTTCCACAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
 1382 CCGTAGGCTGTGAGGAGGCTTCCACAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1441
 1441 AGGCTGAGAGGCTGTGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
 1442 AGGCTGAGAGGCTGTGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1501
 1501 GAGCGCAGCTGT 1560
 1502 GAGCGCAGCTGT 1561
 1561 CACCCAGAGCAACCAATTCAAATTAAGTATGATTCCTCA 1599
 1562 CACCCAGAGCAACCAATTCAAATTAAGTATGATTCCTCA 1600

RESULT 2
 AAH15652
 ID AAH15652 standard; cDNA; 1821 BP.
 XX
 AC AAH15652;

XX	26-JUN-2001	(first entry)
XX	Human cDNA sequence	SEQ ID NO:13996.
DE	Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.	
XX		
XX		
XX		
OS	Homo sapiens.	
XX		
PN	Ep1074617-A2.	
PD	07-FEB-2001.	
XX		
XX	28-JUL-2000; 2000EP-0116126.	
PF		
XX	29-JUL-1999; 99JP-0248036.	
PR	27-AUG-1999; 99JP-0300253.	
PR	11-JAN-2000; 2000JP-0118776.	
PR	02-MAY-2000; 2000JP-0183767.	
PR	09-JUN-2000; 2000JP-0241899.	
XX		
PA	(HELI-) HELIX RES INST.	
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
XX		
XX	WPI; 2001-318749/34.	
XX		
PT	Primer sets for synthesizing polynucleotides, particularly the 5602	
PT	full-length cDNAs defined in the specification, and for the detection	
PT	and/or diagnosis of the abnormality of the proteins encoded by the	
XX	full-length cDNAs -	
XX		
PS	Claim 8; SEQ ID 13996; 2537pp + CD ROM; English.	
XX		
CC	The present invention describes primer sets for synthesizing 5602	
CC	full-length cDNAs defined in the specification. where a primer set	
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary	
CC	to the complementary strand of a polynucleotide which comprises one of	
CC	the 5602 nucleotide sequences defined in the specification, where the	
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination	
CC	of an oligonucleotide comprising a sequence complementary to the	
CC	complementary strand of a polynucleotide which comprises a 5'-end	
CC	sequence and an oligonucleotide comprising a sequence complementary to a	
CC	polynucleotide which comprises a 3'-end sequence, where the	
CC	oligonucleotide comprises at least 15 nucleotides and the combination of	
CC	the 5'-end sequence/3'-end sequence is selected from those defined in	
CC	the specification. The primer sets can be used in antisense therapy and	
CC	in gene therapy. The primers are useful for synthesizing polynucleotides,	
CC	particularly full-length cDNAs. The primers are also useful for the	
CC	detection and/or diagnosis of the abnormality of the proteins encoded by	
CC	the full-length cDNAs. The primers allow obtaining of the full-length	
CC	cDNAs easily without any specialised methods. AAH0156 to AAH13628 and	
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to	
CC	AAB93893 represent human amino acid sequences; and AAH13629 to AAH13632	
CC	represent polynucleotides, all of which are used in the exemplification	
CC	of the present invention.	
XX		
XX		
SO	Sequence 1821 BP; 301 A; 569 C; 614 G; 337 T; 0 other;	
XX		
Query Match	97.6%; Score 1560.8; DB 22; Length 1821;	
Best Local Similarity	98.6%; Pred. No. 0;	
Matches 1574; Conservative	0; Mismatches 22; Indels 0; Gaps 0;	
DB		
3	CCCCGGGGCTCTATAGCCAGAGGGCTCGGGGGGGAAGGCGAGCCCCACAGCGGCGCTG	62
1		
221	CGCGGCTCTCTTACAGCCAGAGGCTCCGGGGGGGAAGGCGAGCCCCACAGCGGCGCTG	280
63	CGAGCGCCCGCTTGGGACACACCGATTAAGAGCTGAAGCAGAGCGCGCCAGCGGACG	122
281	CGAGCGCGCGCTTGGGACACACCGATTAAGAGCTGAAGCAGAGCGCGCCAGCGGACG	340
123	CGCCCGCCACAGCGCGGAGACCCCTGCGAGCGGAGCGCGCGGTGAGATTATGATTC	182

Dp	341	CGCCCCACACGCGCAGGAGCCCCCTGGACAGGAGCCGGGTCGAGGTTATGATCC	400
Qy	183	AGCGGGGCGCCCCCGGGGCTGTCTCCGCGGCCCTGCGCGCTGTGTGTGTGTGAACCC	242
Dp	401	AGCGGGGCGCCCCCGGGGCGGTCTCTCCCGGGCCCTCGCGCGTGTGTGTGTGAACCC	460
Qy	243	GCGGGGGCGCAAGGGCAAGGCCCTTGACAGCTCTTCCGAGTCACTGTCAGCCCCCTTTGGC	302
Dp	461	CGCGGGGCGCAAGGGCAAGGCCCTTGACAGCTCTTCCGAGTCACTGTCAGCCCCCTTTGGC	520
Qy	303	TGAGGCTGAATCTTCCTTACAGCTGATGCTACTAGAGGGGGAAACACGCGCGGAGCT	362
Dp	521	TGAGGCTGAATCTTCCTTACAGCTGATGCTACTAGAGGGGGAAACACGCGCGGAGCT	580
Qy	363	GGTGGGTCGAGAGAGCTGGGCGCGCTGGGAGCGCTCTGGTGTCTCATGTCTGAGACGGCT	422
Dp	581	GGTGGGTCGAGAGAGCTGGGCGCGCTGGGAGCGCTCTGGTGTCTCATGTCTGAGACGGCT	640
Qy	423	GATCAGCAGAGTGTGTAAAGGGCTCTATGAGAGGGCTGTACTGTGGAGACCCGCATCCAGAA	482
Dp	641	GATCAGCAGAGTGTGTAAAGGGCTCTATGAGAGGGCTGTACTGTGGAGACCCGCATCCAGAA	700
Qy	483	GCCCCCTGTGTAGCTCCAGCAGGCTCTGGGCAAGCGGTGGCAGCTTCTTGAACCATTA	542
Dp	701	GCCCCCTGTGTAGCTCCAGCAGGCTCTGGGCAAGCGGTGGCAGCTTCTTGAACCATTA	760
Qy	543	TGCTGGCTATAGCAGAGGTACACCAATGAAGACCTCTGTGACAACTGCAAGCTATTTGCTGTG	602
Dp	761	TGCTGGCTATAGCAGAGGTACACCAATGAAGACCTCTGTGACAACTGCAAGCTATTTGCTGTG	820
Qy	603	CCGCCCGGTGCTTCACCCATGAACCTGCTGTCTGTGACACAGGCTTGGGGGTGCGCTC	662
Dp	821	CCGCCCGGTGCTTCACCCATGAACCTGCTGTCTGTGACACAGGCTTGGGGGTGCGCGCT	880
Qy	663	GTTCTCTGTGCTCAGCCCTGGGCTGGGCTTATGTCTGATGTGGACCTAGAGATGATATA	722
Dp	881	GTTCTCTGTGCTCAGCCCTGGGCTGGGCTTATGTCTGATGTGGACCTAGAGATGATATA	940
Qy	723	GTATCGGCGGTGTGGGGGAGATGACCTTACTCTGGGACCTTCTCGCGTGGAGACCT	782
Dp	941	GTATCGGCGGTGTGGGGGAGATGACCTTACTCTGGGACCTTCTCGCGTGGAGACCT	1000
Qy	783	GGCAGCACTACCGCGGCGCAGCTGGCTACCCCTCCTGTAGAGAGTGGGTTTCAAGACAC	842
Dp	1001	GGCAGCACTACCGCGGCGCAGCTGGCTACCCCTCCTGTAGAGAGTGGGTTTCAAGACAC	1060
Qy	843	TGCTTCCCCCGTTGTGTCTCAGACAGGCGCGGTAGATGACACACTGGTGCCTGAGAGA	902
Dp	1061	TGCTTCCCCCGTTGTGTGTCTCAGACAGGCGCGGTAGATGACACACTGGTGCCTGAGAGA	1120
Qy	903	GCAGGTGCTTCTCTACTGCGAGGTGTGTGTCGCGAGAGACCTTGTGCTAGTCCGGCACT	962
Dp	1121	GCAGGTGCTTCTCTACTGCGAGGTGTGTGTCGCGAGAGACCTTGTGCTAGTCCGGCACT	1180
Qy	963	GCTGACTACGCACTGGCCAGTGTGATGATTTGCTGACACCATGAGGCGCGCTGTGCAAGCTG	1022
Dp	1181	GCTGACTACGCACTGGCCAGTGTGATGATTTGCTGACACCATGAGGCGCGCTGTGCAAGCTG	1240
Qy	1023	CGTCATGCACTGTTTCTACGTGCGGGCGGAGTGTCTGTGCCATGCTGTGCGCTCTT	1082
Dp	1241	CGTCATGCACTGTTTCTACGTGCGGGCGGAGTGTCTGTGCCATGCTGTGCGCTCTT	1300
Qy	1083	CGTGGCCATGAGAGAGGCGAGCAATGAGATGAGAAAGCCCTAATGATTTGATTTGTC	1142
Dp	1301	CGTGGCCATGAGAGAGGCGAGCAATGAGATGAGAAAGCCCTAATGATTTGATTTGTC	1360
Qy	1143	CGTGGTCACCTTCGCTTGGAGCCCAAGAGATGGGAAAGTGTGTTTGCAGTGTGAGGGGA	1202
Dp	1361	CGTGGTCACCTTCGCTTGGAGCCCAAGAGATGGGAAAGTGTGTTTGCAGTGTGAGGGGA	1420
Qy	1203	ATTGATGTTAGCGAGGCCGTGTGACAGGCGCAGGTGCACCCAAATCTTCTGATGTCTAG	1262

Db 1421 ATTGATGTTAGCGAGCCCTGTCAGGCGCCAGGTGACCCAACTACTTCTGATGTCAG 1480
 Oy 1263 CCGTTCGTTGAGACCCCGCCAGCTGAGAGCCCGACAGATGCCACCGCAGAGAGCC 1322
 Db 1481 CCGTTCGTTGAGACCCCGCCAGCTGAGAGCCCGACAGATGCCACCGCAGAGAGCC 1540
 Oy 1323 CTTATGACCCCTGGGCGCGCTGCTGCTAGTGTACTTGTGAGAGACCTCTCTCC 1382
 Db 1541 CTTATGACCCCTGGGCGCGCTGCTGCTAGTGTACTTGTGAGAGACCTCTCTCC 1600
 Oy 1383 CTAAGGCTGAGAGGCTGTCACAGCTCTGTGGGGTGGAGAGACCTCTCTGAGAG 1442
 Db 1601 CTAAGGCTGAGAGGCTGTCACAGCTCTGTGGGGTGGAGAGACCTCTCTGAGAG 1660
 Oy 1443 GGTAGAGAGTGGAGGCTATGCTTGGGGGAGACGCAAGATGATGCTGGGTCAGA 1502
 Db 1661 GGTAGAGAGTGGAGGCTATGCTTGGGGGAGACGCAAGATGATGCTGGGTCAGA 1720
 Oy 1503 GCCAGCTGCTGGGCGCCAGCTGCTATGTAAGGCTCTTGTGTCGAGAGCCCA 1562
 Db 1721 GCCAGCTGCTGGGCGCCAGCTGCTATGTAAGGCTCTTGTGTCGAGAGCCCA 1780
 Oy 1563 CCCACGAGCAATCCAAATGAAGTGAATGCCA 1598
 Db 1781 CCCACGAGCAATCCAAATGAAGTGAATGCCA 1816
 RESULT 3
 AAH16415
 ID AAH16415 standard; cDNA: 1772 BP.
 AAH16415;
 26-JUN-2001 (first entry)
 Human cDNA sequence SEQ ID NO:15393.
 Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 Homo sapiens.
 EP1074617-A2.
 07-FEB-2001.
 28-JUL-2000; 2000EP-0116126.
 29-JUL-1999; 99JP-0248036.
 27-AUG-1999; 99JP-0300253.
 11-JAN-2000; 2000JP-0118776.
 02-MAY-2000; 2000JP-0183767.
 09-JUN-2000; 2000JP-0241899.
 (HELI-) HELIX RES INST.
 Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 WPI: 2001-318749/34.
 Primer sets for synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs -
 Claim 8; SEQ ID 15393; 2537bp + CD ROM; English.
 The present invention describes primer sets for synthesizing 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises at least 15 nucleotides and the combination of
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 CC
 CC
 SQ Sequence 1772 BP: 285 A; 566 C; 585 G; 336 T; 0 other;

Query Match 97.4%; Score 1558.2; DB 22; Length 1772;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 1572; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 3 CCCCCGGGCTCTATAGCAGGCTCCGGGGGAGAGAGCCGCCACAGCCGGCC 62
 Db 178 CCGCGGTCTCTTACAGCCAGGCTCCGGGGGAGAGAGCCGCCACAGCCGGCC 237
 Oy 63 CGAGCCCGCTGGGCGACACCGATTAAGAGCTGAAGAGAGCCGCCACAGCCGGCC 122
 Db 238 CGAGCCCGCTGGGCGACACCGATTAAGAGCTGAAGAGAGCCGCCACAGCCGGCC 297
 Oy 123 CGCCCCACAGCGCCAGGAGCCCTCGGAGGAGAGCCGGGTCAGGTTATGATTC 182
 Db 298 CGCCCCACAGCGCCAGGAGCCCTCGGAGGAGAGCCGGGTCAGGTTATGATTC 357
 Oy 183 AGCGGGGCGCCCGGGGCGTCTCCCGGCGCTCCGGCTGCTGCTGCTGAACCC 242
 Db 358 AGCGGGGCGCCCGGGGCGTCTCCCGGCGCTCCGGCTGCTGCTGCTGAACCC 417
 Oy 243 GCGCGGCGGAG 302
 Db 418 GCGCGGCGGAG 477
 Oy 303 TGAAGGTGAATCTCTTACAGCTGATGCTCACTAGCGGCGGAGAGAGAGAGAG 362
 Db 478 TGAAGGTGAATCTCTTACAGCTGATGCTCACTAGCGGCGGAGAGAGAGAG 537
 Oy 363 GGTGCGGTGAG 422
 Db 538 GGTGCGGTGAG 597
 Oy 423 GATGACAG 482
 Db 598 GATGACAG 657
 Oy 483 GCCCGTGTAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542
 Db 658 GCCCGTGTAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 617
 Oy 543 TGTGCTATATAG 602
 Db 718 TGTGCTATATAG 777
 Oy 603 CCGCCGCTGCTGACAG 662
 Db 778 CCGCCGCTGCTGACAG 837
 Oy 663 GTTCTGTGCTGACAG 722
 Db 838 CTTCTGTGCTGACAG 897
 Oy 723 GTATGGGCTGGGGAG 782

Db 898 GTATGGGCTGTGGGGAGATGCGCTTCACTCTGGGACCTTCTCGGCTGTGGAGCCCT 957
 Oy 783 GCGACCTACCGCGCGGCTGCTACCTCTCTGTAGGAAGATGGTTCGAAGACACC 842
 Db 958 GCGACCTACCGCGCGGCTGCTACCTCTCTGTAGGAAGATGGTTCGAAGACACC 1017
 Oy 843 TGGTTCCTCCGCTGTGGTCCAGAGGCGCGGTAGATGACACCTGGGCGGACGAGGAGA 902
 Db 1018 TGGTTCCTCCGCTGTGGTCCAGAGGCGCGGTAGATGACACCTGGGCGGACGAGGAGA 1077
 Oy 903 GCGAGTGCCTTCTCACTGCGAGTGTGCGCCGAGAGACTTGTGTAGTCTGTGGCACT 962
 Db 1078 GCGAGTGCCTTCTCACTGCGAGTGTGCGCCGAGAGACTTGTGTAGTCTGTGGCACT 1137
 Oy 963 GCTGCACTCGCACTGCGCAATGATGATTTGTGTCACCCATGGGCGGCTGTGCACTGG 1022
 Db 1138 GCTGCACTCGCACTGCGCAATGATGATTTGTGTCACCCATGGGCGGCTGTGCACTGG 1197
 Oy 1023 CGTATGCACTGTTCTTCACTGCGGAGGAGTGTGTCGCACTGCTGCGGCTCTT 1082
 Db 1198 CGTATGCACTGTTCTTCACTGCGGAGGAGTGTGTCGCACTGCTGCGGCTCTT 1257
 Oy 1083 CCGTGCATGAGAGAGGCGAGCATATGAGTATGAAATGCCCTACTTGTATATGTCGC 1142
 Db 1258 CCGTGCATGAGAGAGGCGAGCATATGAGTATGAAATGCCCTACTTGTATATGTCGC 1317
 Oy 1143 CGTGTGCGCTTCCGCTTGTGAGCCCAAGATGGGAAAGTGTGTTGAGTGTGAGGGA 1202
 Db 1318 CGTGTGCGCTTCCGCTTGTGAGCCCAAGATGGGAAAGTGTGTTGAGTGTGAGGGA 1377
 Oy 1203 ATTTGATGTTAGCGAGGCGCGTGTGAGGCGGAGTGTGACCAACTACTTGTGATGTGTAG 1262
 Db 1378 ATTTGATGTTAGCGAGGCGCGTGTGAGGCGGAGTGTGACCAACTACTTGTGATGTGTAG 1437
 Oy 1263 CGGTTGCGTGTGAGGCGCGCGCCAGCTGTGAGAGCCCGCAGAGATGCCACCGCCAGAGAGCC 1322
 Db 1438 CGGTTGCGTGTGAGGCGCGCGCCAGCTGTGAGAGCCCGCAGAGATGCCACCGCCAGAGAGCC 1497
 Oy 1333 CTTATGACCCCTGGGCGCGCTGTGCTTACTTGTCTACTTGTGAGAGCCCTTCTCTTCC 1382
 Db 1498 CTTATGACCCCTGGGCGCGCGCTGTGCTTACTTGTCTACTTGTGAGAGCCCTTCTCTTCC 1557
 Oy 1383 CTAAGGCTGTGAGGCGCTGTGACAGCTCTGTGGGGGTGAGAGAGACTCTGTGAGAGAG 1442
 Db 1558 CTAAGGCTGTGAGGCGCTGTGACAGCTCTGTGGGGGTGAGAGAGACTCTGTGAGAGAG 1617
 Oy 1443 GGTGAGAGAGTGTGAGAGCTATCTTTGGGGGAGACAGCCAGATGAAGTCTGTGGTACAGA 1502
 Db 1618 GGTGAGAGAGTGTGAGAGCTATCTTTGGGGGAGACAGCCAGATGAAGTCTGTGGTACAGA 1677
 Oy 1503 GCCCAGCTGTGGGCGCGGAGCTGCTATGTAAAGGCTTCTAGTTTGTCTGAGAGCCCA 1562
 Db 1678 GCCCAGCTGTGGGCGCGGAGCTGCTATGTAAAGGCTTCTAGTTTGTCTGAGAGCCCA 1737
 Oy 1563 CCCCAGCAACCAATCAAAATGAAGTGAATTTCC 1597
 Db 1738 CCCCAGCAACCAATCAAAATGAAGTGAATTTCC 1772

RESULT 4

AAD04477
 ID AAD04477 standard; CDNA: 1719 BP.

AC AAD04477;

DT 04-JUL-2001 (first entry)

DE Human sphingosine kinase type 1 (hsk1) cDNA.

XX Human: sphingosine kinase type 1; skl; chromosome 17q25.2;
 KW sphingosine-1-phosphate; S1P; drug screening; therapy; hemostasis;
 KW thrombosis; allergic reaction; proliferative disease; cancer;
 KW haematopoietic disorder; leukaemia; cardiovascular disease; stroke;

KW atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
 KW autoimmune disease; inflammatory disease; multiple sclerosis;
 KW T helper-1 related disease; chronic obstructive pulmonary disease;
 KW asthma; myocardial infarction; neurodegenerative disorder;
 KW wound healing; embryogenesis; anticoagulant; cerebroprotective;
 KW neuroprotective; antiproliferative; antitumor; cytoskeletal; cardiac;
 KW vulnerable; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FH CDS 270..1424
 FT /tag= a
 FT /product= "Human sphingosine kinase type 1 (hsk1)."
 FT /note= "CDS is specifically claimed in claim 2 and
 shown as SEQ ID NO 2."
 FT polyA_signal 1675..1681
 FT /tag= b
 FT misc_feature 264..273
 FT /tag= c
 FT /note= "Translational initiator ATG is in a partial
 kozak consensus"

WO200131029-A2.
 PD 03-MAY-2001.
 PD 27-OCT-2000; 2000WO-EP09498.
 PF 28-OCT-1999; 99US-0162307.
 PR 07-FEB-2000; 2000US-0180525.
 PA (WARN) WARNER LAMBERT CO.

PI Allen J, Gosink M, Melendez AJ, Takacs L;
 DR WPI: 2001-300510/31.
 DR P-PSDB: ABE00924.

XX New human sphingosine kinase type I gene for screening drug candidates
 PT particularly inhibitors used for preventing or treating e.g.
 PT atherosclerosis, thrombosis, asthma and diabetes

XX Claim 2: Fig 1: 91pp: English.

XX The present sequence is human sphingosine kinase type 1 (hsk1) cDNA.
 CC The hsk1 gene is located on chromosome 17q25.2. The skl converts the
 CC substrate sphingosine to sphingosine-1-phosphate (S1P). The skl gene
 CC and encoded polypeptide are applicable in screening drug candidates
 CC particularly inhibitors for preventing or treating disorders such as
 CC haemostasis, thrombosis, allergic reactions, proliferative diseases
 CC including cancer, haematopoietic disorders such as leukaemia,
 CC cardiovascular diseases such as stroke, atherosclerosis and coronary
 CC artery disease, dyslipidaemia, diabetes including type I and type II
 CC diabetes, autoimmune and inflammatory diseases such as multiple
 CC sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
 CC disease, asthma, myocardial infarction, neurodegenerative disorders,
 CC natural wound healing processes and embryogenesis.

CC Sequence 1719 BP; 305 A; 529 C; 556 G; 329 T; 0 other;

Query Match 94.8%; Score 1516.2; DB 22; Length 1719;
 Best Local Similarity 97.7%; Pred: No. 4.5e-308;
 Matches 1560; Conservative 0; Mismatches 33; Indels 4; Gaps 2;

Oy 3 CCGCGGCTCTCTATGACGAGCTCCGCGGGGAGAGCCGACACCGGCGCTG 62
 Db 102 CACCGCTCTCTACGACGAGCTCCGCGGGGAGAGAGCCGACACG---CGGCTG 158
 Oy 63 GAGCGCCGCTGTGGGAGAGCTGAAAGAGCTGAAAGAGAGAGAGCGGCGGAG 122
 Db 159 GAGCGCCGCTGTGGGAGAGAGCTGAAAGAGCTGAAAGAGAGAGAGCGGCGGAG 217

PT polynucleotides for diagnosis, prevention and treatment of
PT neurological, cell proliferative and autoimmune/Inflammatory disorders

XX Claim 4; Page 89; 96pp; English.

CC The present sequence encodes a human regulator of intracellular
CC phosphorylation (HRIP). HRIP is useful for screening agonists and
CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
CC are useful for treating a disease or condition associated with
CC decreased or increased expression of functional HRIP. Diseases treated
CC or diagnosed include neurological disorders such as stroke, Parkinson's
CC disease, demyelinating diseases, bacterial and viral meningitis and
CC other developmental disorders of the central nervous system,
CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
CC including leukemia, melanoma, myeloma and cancer of the adrenal gland,
CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
CC inflammatory disorder such as Addison's disease, acquired
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
CC rheumatoid arthritis, microbial infection and trauma.

XX Sequence 1573 BP; 274 A; 480 C; 514 G; 305 T; 0 other:

Query Match 94.5%; Score 1511.6; DB 21; Length 1573;
Best Local Similarity 98.7%; Pred. No. 4.1e-307;
Matches 1534; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 45 GCCCCACAGCCGGCCCTGCGAGCCCGCTGGGCGACACCGATGAAGAGCTGAAGGACG 104
DB 1 GCCCCACAGCCGGCCCTGCGAGCCCGCTGGGCGACACCGATGAAGAGCTGAAGGACG 60
QY 105 AGCCGCCCCGACGGGCGACGCCCCACAGCCCGACAGGACCCCTGCGACGGGAGCCGCG 164
DB 61 AGCCGCCCCGACGGGCGACG-CGCCACAGCCCGACAGGACCCCTGCGACGGGAGCCGCG 119
QY 165 GGTGAGGTATGATGATCCAGCGGGGCGCCCGGGGCGTGTCCCGGAGCCCTGCGGCT 224
DB 120 GGTGAGGTATGATGATCCAGCGGGGCGCCCGGGGCGTGTCCCGGAGCCCTGCGGCT 179
QY 225 GCTGCTGCTGTAACCCCGCGCGGCGGACGAGGCGCTTTCGACGCTTCCGAGTCA 284
DB 180 GCTGCTGCTGTAACCCCGCGCGGCGGACGAGGCGCTTTCGACGCTTCCGAGTCA 239
QY 285 GGTGAGCCCTTTGGCTGAGGCGTGAATGCTTCCGACGCTGAGGCTGACGAGCGCG 344
DB 240 GGTGAGCCCTTTGGCTGAGGCGTGAATGCTTCCGACGCTGAGGCTGACGAGCGCG 299
QY 345 GAACACACGCGGAGCTGTGCGGTGCGAGAGCTGAGGCGCTGGAGCGCTGTGCTGT 404
DB 300 GAACACACGCGGAGGAGCTGTGCGGTGCGAGAGCTGAGGCGCTGGAGCGCTGTGCTGT 359
QY 405 CATGCTGGAAGACGGGCTGATGACAGAGGTGGAAGGGGCTCATGGAAGCGGCGTACTG 464
DB 360 CATGCTGGAAGACGGGCTGATGACAGAGGTGGAAGGGGCTCATGGAAGCGGCGTACTG 419
QY 465 GGAAGACCGCCATCCAGAGCCCGCTGTGAGCTTCCACAGAGCTTGGCAACGCCCTGAC 524
DB 420 GGAAGACCGCCATCCAGAGCCCGCTGTGAGCTTCCACAGAGCTTGGCAACGCCCTGAC 479
QY 525 AGCTTCCTTGAACATATATGCTATGAGTATGAGAGTCAACATGAAGACTTCTGACCA 584
DB 480 AGCTTCCTTGAACATATATGCTATGAGTATGAGAGTCAACATGAAGACTTCTGACCA 539
QY 585 CTGACAGCTATTTGCTGTGCGCGCGGCTGTGACCCATGAACCTGTGCTCTGACAC 644
DB 540 CTGACAGCTATTTGCTGTGCGCGCGGCTGTGACCCATGAACCTGTGCTCTGACAC 599
QY 645 GGCCTCGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 704
DB 600 GGCCTCGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
QY 705 GACCTTAGAGATGATAGTATCGGCGTCTGGGGGAGATGGGCTTCACTGCGGACCTT 764

DB 660 GACCTTAGAGATGATAGTATCGGCGTCTGGGGGAGATGGGCTTCACTGCGGACCTT 719
QY 765 CCTGCTGTGCGACCCCTGCGACCTTACCGGCGCGGAGCTGCTACCTTCCCTGTAGAG 824
DB 720 CCTGCTGTGCGACCCCTGCGACCTTACCGGCGCGGAGCTGCTACCTTCCCTGTAGAG 779
QY 825 AGTGCGTTCAAGACACCTGCTTCCCGGTTGTGTCCAGACGAGGCGCGGTATGATGACA 884
DB 780 AGTGCGTTCAAGACACCTGCTTCCCGGTTGTGTGTCCAGACGAGGCGCGGTATGATGACA 839
QY 885 CCGTGCGCTGAGAGAGAGAGTGCCTTCTCAGTGGAGAGTGTGTGCGGACGAGACTT 944
DB 840 CCGTGCGCTGAGAGAGAGAGTGCCTTCTCAGTGGAGAGTGTGTGCGGACGAGACTT 899
QY 945 TGTGCTATCTGTGCGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
DB 900 TGTGCTATCTGTGCGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
QY 1005 GGGCGCGTGTGAGCTGGGCGTCAATGCTGTTCATGCGGCGGCGGAGTGTCTGTC 1064
DB 960 GGGCGCGTGTGAGCTGGGCGTCAATGCTGTTCATGCGGCGGCGGAGTGTCTGTC 1019
QY 1065 CATGCTGTGCGCTCTTCTCTGCGCATGAGAGAGGCGAGGATATGATGAATGCGC 1124
DB 1020 CATGCTGTGCGCTCTTCTCTGCGCATGAGAGAGGCGAGGATATGATGAATGCGC 1079
QY 1125 CTACTTGTATATGTGCGCGGTGCTGCTTCCGCTTGGAGCCCAAGATGGAAGTGT 1184
DB 1080 CTACTTGTATATGTGCGCGGTGCTGCTTCCGCTTGGAGCCCAAGATGGAAGTGT 1139
QY 1185 GTTGCACTGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1244
DB 1140 GTTGCACTGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
QY 1245 CTACTTGTATATGTGCGCGGTGCTGCTTCCGCTTGGAGCCCAAGATGGAAGTGT 1304
DB 1200 CTACTTGTATATGTGCGCGGTGCTGCTTCCGCTTGGAGCCCAAGATGGAAGTGT 1259
QY 1305 GCCACCGCAGAGAGCCCTTATGACCCCTGCGGCGCGCTGTGCTTATGCTACTTGC 1364
DB 1260 GCCACCGCAGAGAGCCCTTATGACCCCTGCGGCGCGCTGTGCTTATGCTACTTGC 1319
QY 1365 AGGACCTTCT 1424
DB 1320 AGGACCTTCT 1379
QY 1425 GAGACTCTCTGAGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAG 1484
DB 1380 GAGACTCTCTGAGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAG 1439
QY 1485 TGAAGTCTGAGAGAGCCGCTGAGAGCCGCTGAGAGCCGCTGAGAGCCGCTGAGAG 1544
DB 1440 TGAAGTCTGAGAGAGCCGCTGAGAGCCGCTGAGAGCCGCTGAGAGCCGCTGAGAG 1499
QY 1545 TTTGTTGAGAGCCCGCCACCCAGACCAATTAATTAATTAATTAATTAATTAATTAAT 1598
DB 1500 TTTGTTGAGAGCCCGCCACCCAGACCAATTAATTAATTAATTAATTAATTAATTAAT 1553

RESULT 6
AAV84490
ID AAV84490 standard; DNA; 1533 BP.
AAV84490:
01-MAR-1999 (first entry)
XX Human secreted protein gene 80 clone HNF4E54.
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW

KW Immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 XX MO854963-A2.
 XX 10-DEC-1998.
 PD 04-JUN-1998; 98MO-US11422.
 XX 18-DEC-1997; 97US-0070923.
 PR 06-JUN-1997; 97US-0048877.
 PR 06-JUN-1997; 97US-0048881.
 PR 06-JUN-1997; 97US-0048884.
 PR 06-JUN-1997; 97US-0048893.
 PR 06-JUN-1997; 97US-0048896.
 PR 06-JUN-1997; 97US-0048899.
 PR 06-JUN-1997; 97US-0048915.
 PR 06-JUN-1997; 97US-0048949.
 PR 06-JUN-1997; 97US-0048964.
 PR 06-JUN-1997; 97US-0048972.
 PR 06-JUN-1997; 97US-0049020.
 PR 06-JUN-1997; 97US-0049375.
 PR 05-SEP-1997; 97US-0057628.
 PR 05-SEP-1997; 97US-0057635.
 PR 05-SEP-1997; 97US-0057644.
 PR 05-SEP-1997; 97US-0057647.
 PR 05-SEP-1997; 97US-0057650.
 PR 05-SEP-1997; 97US-0057661.
 PR 05-SEP-1997; 97US-0057667.
 PR 05-SEP-1997; 97US-0057761.
 PR 05-SEP-1997; 97US-0057764.
 PR 05-SEP-1997; 97US-0057770.
 PR 05-SEP-1997; 97US-0057775.
 PR 05-SEP-1997; 97US-0057778.
 PR 06-JUN-1997; 97US-0048875.
 PR 06-JUN-1997; 97US-0048878.
 PR 06-JUN-1997; 97US-0048882.
 PR 06-JUN-1997; 97US-0048885.
 PR 06-JUN-1997; 97US-0048894.
 PR 06-JUN-1997; 97US-0048897.
 PR 06-JUN-1997; 97US-0048900.
 PR 06-JUN-1997; 97US-0048910.
 PR 06-JUN-1997; 97US-0048962.
 PR 06-JUN-1997; 97US-0048970.
 PR 06-JUN-1997; 97US-0048974.
 PR 05-SEP-1997; 97US-0049373.
 PR 05-SEP-1997; 97US-0057584.
 PR 05-SEP-1997; 97US-0057629.
 PR 05-SEP-1997; 97US-0057642.
 PR 05-SEP-1997; 97US-0057645.
 PR 05-SEP-1997; 97US-0057648.
 PR 05-SEP-1997; 97US-0057651.
 PR 05-SEP-1997; 97US-0057652.
 PR 05-SEP-1997; 97US-0057668.
 PR 05-SEP-1997; 97US-0057762.
 PR 05-SEP-1997; 97US-0057765.
 PR 05-SEP-1997; 97US-0057771.
 PR 05-SEP-1997; 97US-0057776.
 PR 06-JUN-1997; 97US-0048876.
 PR 06-JUN-1997; 97US-0048880.
 PR 06-JUN-1997; 97US-0048883.
 PR 06-JUN-1997; 97US-0048892.
 PR 06-JUN-1997; 97US-0048895.
 PR 06-JUN-1997; 97US-0048898.
 PR 06-JUN-1997; 97US-0048901.
 PR 06-JUN-1997; 97US-0048917.
 PR 06-JUN-1997; 97US-0048963.
 PR 06-JUN-1997; 97US-0048971.

PR 06-JUN-1997; 97US-0049019.
 PR 06-JUN-1997; 97US-0049374.
 PR 05-SEP-1997; 97US-005627.
 PR 05-SEP-1997; 97US-0057634.
 PR 05-SEP-1997; 97US-0057643.
 PR 05-SEP-1997; 97US-0057646.
 PR 05-SEP-1997; 97US-0057649.
 PR 05-SEP-1997; 97US-0057654.
 PR 05-SEP-1997; 97US-0057666.
 PR 05-SEP-1997; 97US-0057760.
 PR 05-SEP-1997; 97US-0057763.
 PR 05-SEP-1997; 97US-0057769.
 PR 05-SEP-1997; 97US-0057774.
 PR 05-SEP-1997; 97US-0057777.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA;
 PI Fan P, Feng P, Farlie AM, Fischer CL, Florence C;
 PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW;
 PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
 PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
 XX MPI; 1999-059865/05.
 DR P-PSDB; AAW88613.
 XX New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 disorders, immune diseases, inflammation or blood disorders
 XX Claim 4; Page 343; 772pp; English.
 XX The invention relates to nucleic acid sequences (AAW84411 to AAW84633)
 CC encoding human secreted proteins (AAW8534 to AAW8756). The secreted
 CC protein gene sequences are deposited with the ATCC under deposit numbers
 CC AACC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
 CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
 CC cells comprising recombinant vectors containing the nucleic acid
 CC sequences are used for the recombinant production of the secreted
 CC proteins. The polynucleotide and amino acid sequences are useful for are
 CC useful for preventing, treating or ameliorating medical conditions e.g.
 CC by protein or gene therapy. Pathological conditions can be also
 CC diagnosed by determining the presence of mutations in the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the polynucleotides, based on
 CC which tissues they are most highly expressed in, and include developing
 CC products for the diagnosis or treatment of cancer, neurodegenerative
 CC disorders, developmental abnormalities and foetal deficiencies, blood
 CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
 CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
 CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
 CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
 CC such as osteoporosis, arthritis or malignancies, diseases of testes,
 CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The present sequence represents a gene encoding a human secreted protein
 CC (see descriptor line for gene number and clone identification).
 XX
 SQ Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other;
 Query Match 89.2%; Score 1426.8; DB 20; Length 1533;
 Best Local Similarity 97.8%; Pred. No. 2.3e-289;
 Matches 1459; Conservative 6; Mismatches 25; Indels 2; Gaps 2;
 QY 109 GCCGCCACGGGCGAGCCGCCACAGCGCCAGGACCCCTCGGACGAGGGAGCCGGCGGTC 168
 DB 8 GCCGNACAGGGCAGCG-CGCCATAGCCCGCAGGAGCCCGCTGGACAGGGAGCCGGGCTC 66
 QY 169 GAGGTTATGATTCAGAGGGGGGGGGCGGGGGGGTCTCCCGGGCCCTCCCGTACTG 228
 DB 67 GAGGTTATGATTCAGAGGGGGGGGGCGGGGGGGTCTCCCGGGCCCTCCCGTACTG 126
 QY 229 GTGCTGCTGACACCGCGCGGGCGGACAGGCGACGCTTTCGAGACTCACGTG 288

Db	1207	ACGGCAGAAAGGCGCTTATGACACCCCTGGGCGGGCTGTGCTTAGTGTACTATTGCAGG	1266
Oy	1368	ACCCCTCTCTCTTCCCTAAGGGCTCGACAGGCGCTGTTCACAGCTCTGTGTGGGGTGGAGAG	1427
Db	1267	ACCCCTCTCTCTTCCCTAAGGGCTCGACAGGCGCTGTTCACAGCTCTGTGTGGGGTGGAGAG	1326
Oy	1428	ACTCCTCTGGAGAGGGGTGAAGGTGAGGCTTATGCTTTGGGGGGACAGGCCAGGAATGA	1487
Db	1327	ACTCCTCTGGAGAGGGGTGAAGGTGAGGCTTATGCTTTGGGGGGACAGGCCAGGAATGA	1386
Oy	1488	AGTCTGGGTGAGAGAGCCACAGCTGTGGTGGGCCAGTGTCTATGTAAAGGCTTGTAGTTT	1547
Db	1387	AGTCTGGGTGAGAGAGCCACAGCTGTGGTGGGCCAGTGTCTATGTAAAGGCTTGTAGTTT	1446
Oy	1548	GTTCGTGAGACCCCGACCCAGCAAGCAATCAATAAATGAAGTACATTGCCAA	1599
Db	1447	GTTCGTGAGACCCCGACCCAGCAAGCAATCAATAAATGAAGTACATTGCCAA	1498
RESULT 7			
ABAB3273	ID		
ABAB3273	standard; cDNA: 1533 BP		
ABAB3273;			
DT	07-FEB-2002 (first entry)		
DE	Human secreted protein gene 80 SEQ ID NO:90.		
XX			
KW	Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;		
KW	dermatological; immunosuppressive; antineoplastic; immunostimulant;		
KW	cytotoxic; cardiac; vascular; anti-angiogenic; ophthalmological;		
KW	neuroprotective; nontoxic; anticonvulsant; antialzheimers; vulnary;		
KW	antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;		
KW	multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;		
KW	human immunodeficiency virus; hyperproliferative disorder; wound healing;		
KW	Gaucher's disease; cardiovascular disease; Schmitz syndrome; cholestasis;		
KW	Chaga's cardiomyopathy; coronary arteriosclerosis; angioleptic disorder;		
KW	corneal graft neovascularisation; diabetic retinopathy; regeneration;		
KW	neurological disorder; Huntington's chorea; Alzheimer's disease;		
KW	Parkinson's disease; infectious disease; chromosome 17; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200162891-A2.		
XX			
PD	30-AUG-2001.		
XX			
PF	21-FEB-2001; 2001WO-US05614.		
XX			
PR	24-FEB-2000; 2000US-184836P.		
XX			
PR	29-MAR-2000; 2000US-193170P.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
XX	Ni J, Ebnner R, Lafleur DM, Moore PA, Olsen HS, Rosen CA;		
PI	Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;		
PI	Florence C, Hu J, Kyaw H, Fischer CL, Ferris AM, Fan P;		
PI	Peng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;		
PI	Zeng Z, Greene JM;		
XX			
DR	WPI; 2001-625724/72.		
DR	P-PSDB; ABB50380.		
XX			
PT	Nucleic acids encoding 207 human secreted polypeptides, useful for		
PT	preventing, diagnosing and/or treating, e.g. cancers, Parkinson's		
XX	disease and diabetic retinopathy		
XX			
PS	Claim 1; Page 965; 1533pp; English.		
XX			
CC	ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted		
CC	proteins (I) and polynucleotide (II) sequences. (I) and (II) have various		
CC	activities based on the tissues and cells the genes are expressed in.		

CC Example of these activities include: immunomodulatory; antileukemic;
 CC dermatological; immunosuppressive; anti-inflammatory; immunostimulant;
 CC anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;
 CC neuroprotective; nootropic; anticonvulsant; antiallergics; vascular;
 CC antiparkinsonian; antimicrobial; and vulnerary. (i) and (ii) can be used
 CC in gene therapy and vaccine production. (i) and (ii) can be used in the
 CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
 CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
 CC Gaucher's disease), cardiovascular diseases (e.g. scintar syndrome,
 CC Chaga's cardiomyopathy and coronary arteriosclerosis), angogenic
 CC disorders (e.g. corneal graft neovascularisation and diabetic
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
 CC for promoting wound healing, regeneration and/or chemotaxis. ABA8185 to
 CC ABA8193 and ABA8300 represent sequences used in the exemplification of
 CC the present invention.

XX Sequence 1533 BP: 272 A; 454 C; 495 G; 302 T; 10 other:

Query Match 89.2%; Score 1426.8; DB 22; Length 1533;

Best Local Similarity 97.8%; Pred. No. 2.3e-289; Indels 2; Gaps 2;

Matches 1459; Conservative 6; Mismatches 25;

109 GCCGCCACAGGCGACGCCGCCACAGGCGGAGGCCGCCGAGGCGGAGCGGCGGCGT 168
 8 GCCCNCACGGGCGACCG-CCCATAGCGCCAGGAGGCCCGCTGCGAGCGGAGCGGCGGCGT 66
 169 GAGGTATGATGATCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 228
 67 GAGGTATGATGATCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 126
 229 GTGCTGTGAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 288
 127 GTGCTGTGAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 186
 289 CAGCCCGCTTTGGCTGAGGCTGAAATCTCTTCAAGCGTGTACCTGACGAGCGGCGGAGC 348
 187 CAGCCCGCTTTGGCTGAGGCTGAAATCTCTTCAAGCGTGTACCTGACGAGCGGCGGAGC 246
 349 CAGCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 408
 247 CAGCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 306
 409 TCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 467
 307 TTTGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 366
 468 GACCGGCATCCAGAGCCCGCTGTAGCTCCAGCAGGCTGTGCGCAAGCGGCGGCGGCGT 527
 367 GACCGGCATCCAGAGCCCGCTGTAGCTCCAGCAGGCTGTGCGCAAGCGGCGGCGGCGT 426
 528 TTCCCTTGAACATATGCTGCTATGAGAGGTCACCAATGAAGCCTCGACCACTG 587
 427 TTCCCTTGAACATATGCTGCTATGAGAGGTCACCAATGAAGCCTCGACCACTG 486
 588 CAGGCTATGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 647
 487 CAGGCTATGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 546
 648 TTGCGGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707
 547 TTGCGGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
 708 CCAAGAGAGTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
 607 CCAAGAGAGTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
 768 GCGCTGAGAGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 827
 667 GCGCTGAGAGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 726
 828 GGGTTTCAAGACACTGCTTCCCGCTTGTGTGTCAGAGGCGGCGGCTGATGACACCT 887

DB 727 GGGTTTCAAGACACTGCTTCCCGCTTGTGTGTCAGAGGCGGCGGCTGATGACACCT 786
 QY 888 GGTCCACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 947
 DB 787 TGTCCACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 846
 QY 948 GGTCCACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1007
 DB 847 GGTCCACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
 QY 1008 CCGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1067
 DB 907 CCGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
 QY 1068 GCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1127
 DB 967 GCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1026
 QY 1128 CTTGATATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
 DB 1027 CTTGATATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086
 QY 1188 TGCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1247
 DB 1087 TGCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1146
 QY 1248 CTTGATATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1307
 DB 1147 CTTGATATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206
 QY 1308 ACCGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1367
 DB 1207 ACCGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266
 QY 1368 ACCCTTCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1427
 DB 1267 ACCCTTCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1326
 QY 1428 ACTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1487
 DB 1327 ACTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1386
 QY 1488 AGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1547
 DB 1387 AGTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
 QY 1548 GTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1599
 DB 1447 GTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1498

RESULT 8

AA159336 standard; cDNA: 1438 BP.

AA159336:

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 1539.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokine; thrombolytic; drug screening; arthritis; inflammation;
 leukaemia; ss.

Homo sapiens.

WO200153312-A1.

PD 26-JUL-2001.
XX 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao Q, Zhou P, Goodrich R, Dirmannac RT;
XX
DR WPI: 2001-442253/47.
P-PSDB: AAM40180.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1: SEQ ID NO 1539; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilization of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 1438 BP; 251 A; 428 C; 463 G; 296 T; 0 other:
SQ

Query Match 87.1%; Score 1393; DB 22; Length 1438;
Best Local Similarity 98.6%; Pred. No. 2,7e-282;
Matches 1405; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 535 AACATTAAGCTGGCTATGAGCAGGTCACCAATGAAGACCTCTGACCAACTGCACGGTA 594
DB 361 AACCATTAAGCTGGCTATGAGCAGGTCACCAATGAAGACCTCTGACCAACTGCACGGTA 420
QY 595 TTGCTGTGCGCGCCGGCTGTCACCAATGAAGACCTCTGTCACACAGGCTTGGGG 654
DB 421 TTGCTGTGCGCGCGCTGTCACCAATGAAGACCTCTGTCACACAGGCTTGGGG 480
QY 655 CTGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
DB 481 CTGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 715 AGTGAATGAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774
DB 541 AGTGAATGAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 775 GCAGCCCTGCGACCTACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
DB 601 GCAGCCCTGCGACCTACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 835 AAGACACTGCTTCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
DB 661 AAGACACTGCTTCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 895 CTGAGGAGCAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
DB 721 CTGAGGAGCAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 955 CTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
DB 781 CTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 1015 GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074
DB 841 GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1075 GCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1134
DB 901 GCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 1135 TATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1194
DB 961 TATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1195 GATGGGAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1254
DB 1021 GATGGGAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1080
QY 1255 ATGTGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1314
DB 1081 ATGTGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1315 GAAGACCTTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1374
DB 1141 GAAGACCTTATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1200
QY 1375 CTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1434
DB 1201 CTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY 1435 TGGAGAAGGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1494
DB 1261 TGGAGAAGGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1320
QY 1495 GGTCAAGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1554
DB 1321 GGTCAAGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
QY 1555 GACCCCAACCCACAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1599
DB 1381 GACCCCAACCCACAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1425

RESULT 9
ID AAA50508 standard; cDNA; 1447 BP.
XX
AC AAA50508;
XX
DT 05-DEC-2000 (first entry)
XX
DE Human sphingosine kinase A cDNA.
XX
KW Sphingosine kinase A; SKA; human; drug screening; infection;
KW anti-inflammatory; anti-allergic; anticancer; inflammation; allergy;
KW cancer; therapy; diagnosis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 7..1161
FT /*tag= a
XX
XX MO200052173-A2.
XX
XX 08-SEP-2000.
XX
XX 02-MAR-2000; 2000MO-CA00223.
XX
XX 02-MAR-1999; 99US-0122516.
XX
XX (ALIX) NPS ALLELIX CORP.
XX
XX Munroe D, Gupta A, Falzone GR;
XX
XX WPI: 2000-572185/53.
XX
XX P-PSDB: AAY06057.
XX
XX New human sphingosine kinase A, B and C polynucleotides and
XX polypeptides useful in e.g. chromosome and gene mapping, and detecting
XX inflammation or disease associated with abnormal levels of sphingosine
XX kinase expression
XX
XX
XX Disclosure: Fig 1; 81pp; English.
XX
XX The present sequence is that of an isolated polynucleotide encoding
XX human sphingosine kinase A (SKA, see AAY96057), an enzyme that
XX phosphorylates sphingosine to form sphingosine 1-phosphate.
XX The polynucleotide was isolated from an Hela cDNA library by
XX PCR amplification. The invention provides polynucleotides (see
XX AAA50508-10) and polypeptides (see AAY96057-59) for the human
XX sphingosine kinase (SK) homologues SKA, SKB and SKC. The
XX polynucleotides may be used as hybridization probes, in the
XX construction of PCR primers for chromosome and gene mapping, in
XX the recombinant production of SKA, SKB and SKC, and in the
XX generation of antisense DNA or RNA. They can be used to detect
XX inflammation or disease associated with abnormal levels of SK
XX expression, or to detect differences in gene sequence between
XX normal and carrier or affected individuals. Host cells expressing
XX SK can be used in drug screening. Human SK specific antibodies,
XX inhibitors, ligands or their analogues can be used as bioactive
XX agents to treat inflammation or disease including viral, bacterial
XX or fungal infections, allergic responses, mechanical injury
XX associated with trauma, hereditary diseases, lymphoma or carcinoma,
XX and other conditions with activate the genes of kidney, lung,
XX heart, lymphoid or tissues of the nervous system.
XX
XX Sequence 1447 BP; 256 A; 427 C; 465 G; 299 T; 0 other;
SQ
Query Match 87.0%; Score 1391; DB 21; Length 1447;
Best Local Similarity 98.3%; Pred. No. 7.1e-282;
Matches 1406; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 169 GAGGTTATGATTCAGCGGGGGCCCCCGGGGGGTCTCCCGGGGGCCCTGCGGCGTCTG 228
|||||

Db 1 GAGGTTATGATTCAGCGGGGGCCCCCGGGGGGTGTCTCCCGGGGGCCCTGCGGCGTCTG 60
QY 229 GTGCTGCTGAACCCGCGGGGGGCAAGGCAAGCCCTTGAGCTCTTCCGGAGTCACTG 288
Db 61 GTGCTGCTGAACCCGCGGGGGGCAAGGCAAGCCCTTGAGCTCTTCCGGAGTCACTG 120
QY 289 CAGCCCTTTGGGTGAGGCTGAATCTCTTCACGCTGATGCTCACTAGCGCGGCAAC 348
Db 121 CAGCCCTTTGGGTGAGGCTGAATCTCTTCACGCTGATGCTCACTAGCGCGGCAAC 180
QY 349 CAGCGCGGGAGCTGTGCGGTGAGAGAGCTGGGCGCTGGAGCTCTGTGTGTCTATG 408
Db 181 CAGCGCGGGAGCTGTGCGGTGAGAGAGCTGGGCGCTGGAGCTCTGTGTGTCTATG 240
QY 409 TCTGGAGCGGGCTGTATGACAGAGTGTGAAGGGGCTCAATGAGCGGCTGACTGGAG 468
Db 241 TCTGGAGCGGGCTGTATGACAGAGTGTGAAGGGGCTCAATGAGCGGCTGACTGGAG 300
QY 469 ACCGCCATCCAGAAACCCCTGTGTAGCTCCAGCAGAGCTCTGCAACGCGCTGGCAGCT 528
Db 301 ACCGCCATCCAGAAACCCCTGTGTAGCTCCAGCAGAGCTCTGCAACGCGCTGGCAGCT 360
QY 529 TCTTGAACCATTTATGCTGTGCTATGAGAGGTCACCATGAAAGCTCTGACCACTGC 588
Db 361 TCTTGAACCATTTATGCTGTGCTATGAGAGGTCACCATGAAAGCTCTGACCACTGC 420
QY 589 ACGCTATTGCTGTGCGCGCGGTGCTGTACCATGAACCTGTGTCTGTGACAGGCT 648
Db 421 ACGCTATTGCTGTGCGCGCGGTGCTGTACCATGAACCTGTGTCTGTGACAGGCT 480
QY 649 TCGGGGCTGCGCTGTTCTGTGTGTCAAGCTTGCGCTGGGCTTCATGTGTGTGAC 708
Db 481 TCGGGGCTGCGCTGTTCTGTGTGTCAAGCTTGCGCTGGGCTTCATGTGTGTGAC 540
QY 709 CTGAGAGATGATTAAGATGAGGGGTGGGGGAGATGCGCTTCACTGTGGGACCTTCTG 768
Db 541 CTGAGAGATGATTAAGATGAGGGGTGGGGGAGATGCGCTTCACTGTGGGACCTTCTG 600
QY 769 CGTGTGGCAGCCCTTGCGCACTACCGCGGCGAGCTGCTACCTCTCTGTAGAAAGATG 828
Db 601 CGTGTGGCAGCCCTTGCGCACTACCGCGGCGAGCTGCTACCTCTCTGTAGAAAGATG 660
QY 829 GGTTCAGACACCTCTCCCTGTTGTGTGTCAGCAGAGGCGGGTATGATCACACCTG 888
Db 661 GGTTCAGACACCTCTCCCTGTTGTGTGTCAGCAGAGGCGGGTATGATCACACCTT 720
QY 889 GTGCCACTGGAGAGCAGGTGCTTCTCACTGTGACAGTGTGTGCGGAGTCTGTGCAATG 948
Db 721 GTGCCACTGGAGAGCAGGTGCTTCTCACTGTGACAGTGTGTGCGGAGTCTGTGCAATG 780
QY 949 CTAGCTCTGCACTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 1008
Db 781 CTAGCTCTGCACTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 840
QY 1009 CGCTGTGCACTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 1068
Db 841 CGCTGTGCACTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 900
QY 1069 CTGCTGCGCTCTTCTGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
Db 901 CTGCTGCGCTCTTCTGCGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 1129 TTGCTATATGCGCCCGGTGCGCTTCCGTTGAGAGCCCAAGAGATGGAAGGTGTGTT 1188
Db 961 TTGCTATATGCGCCCGGTGCGCTTCCGTTGAGAGCCCAAGAGATGGAAGGTGTGTT 1020
QY 1189 GCAGTGAATGGGGAATGATGTTAGCAGAGCCGTGCAAGGCGCAAGTGTCAACCAACTAC 1248
Db 1021 GCAGTGAATGGGGAATGATGTTAGCAGAGCCGTGCAAGGCGCAAGTGTCAACCAACTAC 1080
QY 1249 TTCTGATGTGACGGGTGTGCTGTGAGAGCCCGCGCGAGTGAAGCCCGCAGCATGCCA 1308
Db 1081 TTCTGATGTGACGGGTGTGCTGTGAGAGCCCGCGCGAGTGAAGCCCGCAGCATGCCA 1140

[illegible]

DR P-PSDB: AAB48007 .
 XX Novel sphingosine kinase protein and nucleic acid molecules for
 PT diagnosis, prophylaxis and treatment of rheumatoid arthritis, asthma,
 PT atherosclerosis, inflammation, meningitis, multiple sclerosis and
 PT septic shock -
 XX
 XX
 XX Claim 4; Fig 7a; 100pp: English.
 XX
 XX This CDNA encodes a human sphingosine kinase (SK) protein. The human SK
 CC protein, encoding nucleic acids and modulators are useful for modulating
 CC expression, functional activity or cellular functional activity of
 CC sphingosine kinase in a subject and also for treating a mammal by
 CC modulating the activity of SK. Diseases treated by regulating SK
 CC cellular activity include rheumatoid arthritis, asthma, atherosclerosis,
 CC inflammation, meningitis, multiple sclerosis and septic shock.
 XX
 XX Sequence 1205 BP; 196 A; 357 C; 400 G; 252 T; 0 other;
 SQ

Query Match	71.48;	Score 1141.4;	DB 22;	Length 1205;
Best Local Similarity	97.88;	Pred. NO. 1.2e-229;		
Matches 1157; Conservative	0;	Mismatches 26;	Indels 0;	Gaps 0;

QY	149	GGCAGCGGAGACCCCGGGGTGCAGAGTTATGATATCCAGCGGCGGCCCCCGGGGCTGTCTCC	208
Db	7	GGCAGAGAGAGCGCGGGGTGAGGTTATGATATCCAGCGGCGGCCCCCGGGGCTGTCTCC	66
QY	209	CGCGGCGCCGACCGCGCTGTGTGTCTGTCTGTAAACCGCGCGCGCGGCGCAAGGCGACCTTGC	268
Db	67	CGCGGCGCCGACCGCGCTGTGTGTCTGTCTGTAAACCGCGCGCGCGGCGCAAGGCGACCTTGC	126
QY	269	AGCTCTTCCGAGATACAGTGCAGCGCCCTTTTGGGCTGAGGCTGAAATCTCTTCACGCTGA	328
Db	127	AGCTCTTCCGAGATACAGTGCAGCGCCCTTTTGGGCTGAGGCTGAAATCTCTTCACGCTGA	186
QY	329	TGCTCACTAGCGGCGGAGAACACACGCGCGGAGCTGTGCGGTGCGAGGACGCTGGGCGCT	388
Db	187	TGCTCACTAGCGGCGGAGAACACACGCGCGGAGCTGTGCGGTGCGAGGAGCTGGGCGCT	246
QY	389	GAGAGCCTCTGGTGTCTATGTCTGGAGACGGGCTATATGACAGAGGTGGGAAACGGGCTCA	448
Db	247	GAGAGCCTCTGGTGTCTATGTCTGGAGAGGGGCTATATGACAGAGGTGGGAAACGGGCTCA	306
QY	449	TGAGAGGGGCTGACTGTGGAGACCGGCATCCAGAAAGCCCTGTGTAGCCCTCCACGAGCT	508
Db	307	TGAGAGGGGCTGACTGTGGAGAGCCGCATCCAGAAAGCCCTGTGTAGCCCTCCACGAGCT	366
QY	509	CTGGCAACGGCGTGGAGACTTCTTGAACCATTATCTGTGGCTATAGCAGAGTCCACCAATG	568
Db	367	CTGGCAACGGCGTGGAGACTTCTTGAACCATTATCTGTGGCTATAGCAGAGTCCACCAATG	426
QY	569	AAGACCTCTGACCAATGCAAGCTATTGTCTGCGCGCCGGGTCGTACCCATGAAAC	628
Db	427	AAGACCTCTGACCAATGCAAGCTATTGTCTGCGCGCCGGGTCGTACCCATGAAAC	486
QY	629	TGCTGTCTGCAACACGGCTTCGGGGCTGCGCTGTCTGTGTGCTCAAGCTGCGCTGGG	688
Db	487	TGCTGTCTGCAACACGGCTTCGGGGCTGCGCTGTCTGTGTGCTCAAGCTGCGCTGGG	546
QY	689	GCTTCATTGCTGATGTGACCTAAGAGATATAGTTCGGGCTCTGCGGGAATGACGCT	748
Db	547	GCTTCATTGCTGATGTGACCTAAGAGATATAGTTCGGGCTCTGCGGGAATGACGCT	606
QY	749	TCACGTGGGACCTTCTCGTGTGACAGGCTGCGACCTACCGCGGCGGACGTGCTA	808
Db	607	TCACGTGGGACCTTCTCGTGTGACAGGCTGCGACCTACCGCGGCGGACGTGCTA	666
QY	809	CCCTCCCTGTAGGAAGATGGGTTTCAAGACACTGCTTCCCGGTTGTGTCCAGGAGG	868
Db	667	ACCTCCCTGTAGGAAGATGGGTTTCAAGACACTGCTTCCCGGTTGTGTGTCCAGGAGG	726
QY	869	GCCCGGTAGTGCACACTGGTGGCAGCTGGAGAGAGAGTGCCTTCTCACTGCGAGGTGG	928
Db			

Db 727 GCCCGTAGATGACACACCTGTGCGCATGAGAGACCAAGTGCCTCTCATCTGACATGG 786
Qy 929 TGCCGACGAGACTTGTGTAGTCTGAGCAGTCTGACCTGACCTGCGGAGTGA 988
Db 787 TGCCGACGAGACTTGTGTAGTCTGAGCAGTCTGACCTGACCTGCGGAGTGA 846
Qy 989 TGTGTGTCACACCAATGGGCGCGTGTGACAGTGGCGTCAATCTGTTTACGTGCGG 1048
Db 847 TGTGTGTCACACCAATGGGCGCGTGTGACAGTGGCGTCAATCTGTTTACGTGCGG 906
Qy 1049 CGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1108
Db 907 CGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
Qy 1109 TGAATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1168
Db 967 TGAATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
Qy 1169 AGGATGGGAAAGT 1228
Db 1027 AGGATGGGAAAGT 1086
Qy 1229 GCCAGGTGACCCCAACTACTTCTGATGCTGACGCTTGGTGAGCCCGCCAGCT 1288
Db 1087 GCCAGGTGACCCCAACTACTTCTGATGCTGACGCTTGGTGAGCCCGCCAGCT 1146
Qy 1289 GGAAGCCCGACAGATGCGCCGCGGAGAGCCCTTATGACC 1331
Db 1147 GGAAGCCCGACAGATGCGCCGCGGAGAGCCCTTATGACC 1189

RESULT 12

ABL59532
ID ABL59532 standard; cDNA; 1173 BP.

AC ABL59532;

DT 16-JUL-2002 (first entry)

DE Human sphingosine kinase (SphK1) cDNA SEQ ID NO:32.

KM Human; sphingosine kinase; SphK1; enzyme; chromosome 17q25.2; gene;
tumour; lipid associated gene; lipid metabolism; lipid synthesis; ss.

OS Homo sapiens.

PN W0200227028-A1.

PD 04-APR-2002.

PF 27-SEP-2001; 2001MO-US30366.

PR 28-SEP-2000; 2000US-0676052.

PA (ATAI-) ATAIRGIN TECHNOLOGIES INC.

PI Skinner MK, Patton JL, Chaudhary J;

DR WPI: 2002-402054/43.

PT Identifying tumor characteristics in a tissue sample taken from a
patient, involves determining the copy number or expression level of
genes associated with lipid metabolism, synthesis or action

PS Example 1; Page 90; 113pp; English.

CC The present invention describes a method for identifying tumour
characteristics, comprising measuring a copy number or expression level
of at least two genes associated with lipid metabolism, synthesis, or
action in cells from a patient tissue sample, and comparing the results
with a copy number or expression level of the genes in a normal cell.
CC Also described is an array of nucleic acid polymers immobilised on a
solid support, comprising a solid support, at least two different nucleic

CC acid polymers which are each specific for a different gene associated
with lipid metabolism, synthesis or action, where each nucleic acid
polymer is located at a predetermined position on the solid support, and
the array comprises nucleic acid polymers which are specific for less
than 100 genes other than the selected genes. The method is useful for
determining tumour characteristics in a tissue sample taken from a
patient. The present sequence represents a human lipid-associated gene
related cDNA sequence, which is used in the exemplification of the
present invention.

XX Sequence 1173 BP; 188 A; 349 C; 390 G; 246 T; 0 other;

Query Match 71.2%; Score 1137.8; DB 24; Length 1173;

Best Local Similarity 98.1%; Pred. No. 6; 7e-229;

Matches 1151; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 157 GAGCCGCGGCTGAGGTTATGATTCACGCGGCGCCCGCGGCGTCCCGGCGCC 216
Db 1 GAGCCGCGGCTGAGGTTATGATTCACGCGGCGCCCGCGGCGTCCCGGCGCC 60
Qy 217 TGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
Db 61 TGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 277 CGGAGTCACTGACACCCCTTTTGGCTGAGGCTGAATCTCTTCACTGATGCTCACT 336
Db 121 CGGAGTCACTGACACCCCTTTTGGCTGAGGCTGAATCTCTTCACTGATGCTCACT 180
Qy 337 GAGCGCGGAAACACCGCGGAGCTGTGCGGTGAGAGAGTGGCCGCTGGAGAGCT 396
Db 181 GAGCGCGGAAACACCGCGGAGCTGTGCGGTGAGAGAGTGGCCGCTGGAGAGCT 240
Qy 397 CTGCTGTGCTGTGAGAGAGCGGCTGATGACGAGAGGTGAACGGGCTCATGAGAGG 456
Db 241 CTGCTGTGCTGTGAGAGAGCGGCTGATGACGAGAGGTGAACGGGCTCATGAGAGG 300
Qy 457 CCTGACTGGAGACCGCCATCCAGAGCCCTGTGTAGCTTCCAGAGGCTTGTGCAAC 516
Db 301 CCTGACTGGAGACCGCCATCCAGAGCCCTGTGTAGCTTCCAGAGGCTTGTGCAAC 360
Qy 517 GCGCTGGACCTTCTTGAACCTTATGCTGTGCTGTGAGAGGTACCAATGAGAGCTC 576
Db 361 GCGCTGGACCTTCTTGAACCTTATGCTGTGCTGTGAGAGGTACCAATGAGAGCTC 420
Qy 577 CTGACCACTGACGCTATGCTGTGCGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCT 636
Db 421 CTGACCACTGACGCTATGCTGTGCGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCT 480
Qy 637 CTGACACGCTTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db 481 CTGACACGCTTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 697 GCTGATGTGACCTAGAGATGATAGTATGCGGCTGTGAGGAGATGCGCTTCACTCTG 756
Db 541 GCTGATGTGACCTAGAGATGATAGTATGCGGCTGTGAGGAGATGCGCTTCACTCTG 600
Qy 757 GGCACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
Db 601 GGCACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 817 GTAGAGAGTGGGTTTCAAGACACCTGCTTCCCGCTTGTGTGCTGCTGCTGCTGCTGCT 876
Db 661 GTAGAGAGTGGGTTTCAAGACACCTGCTTCCCGCTTGTGTGCTGCTGCTGCTGCTGCT 720
Qy 877 GATGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 936
Db 721 GATGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 937 GAGGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
Db 781 GAGGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 997 GCACCATGAGCGCGCTGTGACACTGCGCTCATGATCTGTTCTTACGTGCGGCGGAGTGG 1056

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 12:35:23 ; Search time 4153 seconds

(without alignments)
11205.246 Million cell updates/sec

Title: US-09-784-810a-1_COPY_2_1600

Perfect score: 1599
Sequence: 1 cccccggggtctctatagc.....aataaagtgacattcccaa 1599Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Genembl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1600	6	AX224379 Sequence
2	1560.8	97.6	1821	9	AK022402 Homo sapi
3	1558.2	97.4	1772	9	AK023393 Homo sapi
4	1557	97.4	1824	9	BC008040 Homo sapi
5	1549.6	96.9	1783	9	AF238083 Homo sapi
6	1549.6	94.8	1719	6	AX127641 Sequence
7	1516.2	94.8	1693	9	AF267566 Homo sapi
8	1505.6	94.2	1869	9	BC009419 Homo sapi
9	1385	86.6	2502	9	AK095578 Homo sapi
10	1342.2	83.9	1428	9	HS245504 Homo sapi
11	1314.8	82.2	2015	9	AB046025 Macaca fa
12	1137.8	71.2	1173	9	AF200328 Homo sapi
13	1124.6	70.3	1155	6	AX127642 Sequence
14	1103	69.0	1192	9	BC004112 Homo sapi
15	1017.4	63.6	18343	2	AC068145 Homo sapi
16	1008.6	63.1	209861	2	AC021196 Homo sapi
17	836.8	52.3	1815	6	AX287137 Sequence
18	836.8	52.3	1815	10	AF068748 Mus muscu
19	836.4	52.3	1759	6	AX224381 Sequence
20	818.4	51.2	1842	10	AB049573 Rattus no
21	818.4	51.2	1895	10	AB049572 Rattus no
22	818.4	51.2	1943	10	AB049571 Rattus no
23	818.4	51.2	2501	10	AB049574 Rattus no
24	818.4	51.2	2648	10	AB049575 Rattus no
25	769.6	48.1	1559	6	AX287138 Sequence
26	769.6	48.1	1559	10	AF068749 Mus muscu
27	737.6	46.1	1146	10	AF415213 Mus muscu
28	513.6	32.1	110000	2	AL606505_0
29	513.6	32.1	112037	10	AL645851 Mouse DNA
30	382.6	23.9	394	6	AX334274 Sequence
31	258	16.1	2380	9	AF245447 Homo sapi
32	258	16.1	2731	9	BC010671 Homo sapi
33	258	16.1	2875	6	AX086301 Sequence
34	258	16.1	2875	9	HS2801669 Homo sapi
35	245.2	15.3	3012	9	BC006161 Homo sapi
36	245.2	15.3	1854	10	AF415214 Mus muscu
37	245.2	15.3	2416	10	BC006941 Mus muscu
38	242	15.1	2698	10	AF245448 Mus muscu
39	240	15.0	240	6	AX127662 Sequence
40	193	12.1	296	6	AX198252 Sequence
41	193	12.1	296	6	AX208818 Sequence
42	177.2	11.1	62875	2	AC068436 Homo sapi
43	173.6	10.9	44767	2	AC106591 Rattus no
44	173	10.8	254729	2	AC122611 Rattus no
45	159.2	10.0	1172	9	AK000599 Homo sapi

ALIGNMENTS

RESULT 1	AX224379	1600 bp	DNA	linear	PAT 10-SBP-2001
LOCUS	AX224379				
DEFINITION	Sequence 1 from Patent WO0160990.				
ACCESSION	AX224379				
VERSION	AX224379.1	GI:15554631			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	1 (bases 1 to 1600)				
JOURNAL	Novel sphingosine kinases				
	Rastelli, L.				
	Patent: WO 0160990-A 1 23-AUG-2001;				

Curagen Corporation (US) : GENENTECH, INC. (US)

FEATURES Location/Qualifiers

Source 1..1600 /organism="Homo sapiens" /db_xref="taxon:9606"

BASE COUNT 265 a 492 c 531 g 311 t 1 others

ORIGIN

Query Match 100.0%; Score 1599; DB 6; Length 1600; Best Local Similarity 100.0%; Pred. No. 4,8e-272; Matches 1599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCCCGGGGCTCTATAGCCAGGCTCCGGGCGGGGAAGGCCAGGCCACAGCCGCC 60
 2 CCCCCGGGGCTCTATAGCCAGGCTCCGGGCGGGGAAGGCCAGGCCACAGCCGCC 61
 61 TCGCAGCCGCCCTTGGGAGCAGCAGATAGAGAGTGAAGGAGGAGCGGCCAGGAG 120
 62 TCGCAGCCGCCCTTGGGAGCAGCAGATAGAGAGTGAAGGAGGAGCGGCCAGGAG 121
 121 AGCGCCGCCAGCGCAGGAGCCCTGGCAGCGGAGCGCGGGTCTGAGTTATGAT 180
 122 AGCGCCGCCAGCGCAGGAGCCCTGGCAGCGGAGCGCGGGTCTGAGTTATGAT 181
 181 CCAGCGGGGCGCCCGGGGCGCTCTCCCGGGCCCTGCGCTGCTGCTGCTGAC 240
 182 CCAGCGGGGCGCCCGGGGCGCTCTCCCGGGCCCTGCGCTGCTGCTGCTGAC 241
 241 CCGCGGGGCGGCAAGGCGCTTGGAGCTCTTCCGAGTCACTGTCAGCCCTTTTG 300
 242 CCGCGGGGCGGCAAGGCGCTTGGAGCTCTTCCGAGTCACTGTCAGCCCTTTTG 301
 301 GCTGAGGCTGAATCTCTTCACTGATGCTCATCTGAGCGCGGAGAACAGCGGGAG 360
 302 GCTGAGGCTGAATCTCTTCACTGATGCTCATCTGAGCGCGGAGAACAGCGGGAG 361
 361 CTGAGGCTGAGAGAGCTGGGCGCGCTGGAGCGCTGCTGCTGCTGCTGAGACGAG 420
 362 CTGAGGCTGAGAGAGCTGGGCGCGCTGGAGCGCTGCTGCTGCTGCTGAGACGAG 421
 421 CTGAGTCAAGAGTGTGAGAGCGGCTCATGAGAGCGGCTGACTGGGAGACCGCCATCCAG 480
 422 CTGATCAAGAGTGTGAGAGCGGCTCATGAGAGCGGCTGACTGGGAGACCGCCATCCAG 481
 481 AAGCCCTGTGTAGCTCCAGAGAGCTGTGGCAAGCGCTGCGCACTCTTGAACAT 540
 482 AAGCCCTGTGTAGCTCCAGAGAGCTGTGGCAAGCGCTGCGCACTCTTGAACAT 541
 541 TATGCTGGCTATGAGAGGTACCAATGAAGACCTCTGACCACTGACAGCTATTGCTG 600
 542 TATGCTGGCTATGAGAGGTACCAATGAAGACCTCTGACCACTGACAGCTATTGCTG 601
 601 TCGCGCGGGTGTCTTACACCATAGAACCTGCTGCTGAGACAGGCTTCGGGGCTGCCG 660
 602 TCGCGCGGGTGTCTTACACCATAGAACCTGCTGCTGAGACAGGCTTCGGGGCTGCCG 661
 661 TCGTCTCTGTGTCAAGCTGTGCGCTGGGCTTCAATTGCTGATGTGAGACCTAGAGATGAT 720
 662 TCGTCTCTGTGTCAAGCTGTGCGCTGGGCTTCAATTGCTGATGTGAGACCTAGAGATGAT 721
 721 AAGTATGGGCTGTGGGGAGAGATGCGCTTACACCTTGGGCACTTCCTGGGCTGGAGCC 780
 722 AAGTATGGGCTGTGGGGAGAGATGCGCTTACACCTTGGGCACTTCCTGGGCTGGAGCC 781
 781 CTGGGCACTTACCGCGGCGAGCTAGCTACCTCCCTTAGAGAGAGTGGCTTCAAGACA 840
 782 CTGGGCACTTACCGCGGCGAGCTAGCTACCTCCCTTAGAGAGAGTGGCTTCAAGACA 841
 841 CCTGCTTCCCGGTTGGTTCAGACAGGCGCGGTATAGTACACCTTGGTCCACTGAG 900
 842 CCTGCTTCCCGGTTGGTTCAGACAGGCGCGGTATAGTACACCTTGGTCCACTGAG 901
 901 GAGCAGGTGCTTCTCACTGGCAGGTGGTGGCCAGAGGACTTGTGCTAGTCTTGCA 960

|||||

Db 902 GAGCAGGTGCTTCTCACTGCGAGGTGCTGCCAGAGAGACTTGTGCTAGTCTGCA 961

|||

Qy 961 CTGCTGCACTGCGACCTGGCCAGTGAATGTTTGTCTGACACCAATGGCGCTGCGAGCT 1020

Db 962 CTGCTGCACTGCGACCTGGCCAGTGAATGTTTGTCTGACACCAATGGCGCTGCGAGCT 1021

|||

Qy 1021 GCGCTCATCATCTGTTCTGAGTGGGGGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCT 1080

Db 1022 GCGCTCATCATCTGTTCTGAGTGGGGGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCT 1081

|||

Qy 1081 TTCTTGGCCATGAGAGAGGCGAGCATATGAGTGAATGCAATGCTTCTGCTGCTGCTGCT 1140

Db 1082 TTCTTGGCCATGAGAGAGGCGAGCATATGAGTGAATGCAATGCTTCTGCTGCTGCTGCT 1141

|||

Qy 1141 CCGCTGCTGCTTCTGCTGAGAGCCCAAGAGTGGGAAAGTGTGCTTGGAGTGAATGAGG 1200

Db 1142 CCGCTGCTGCTTCTGCTGAGAGCCCAAGAGTGGGAAAGTGTGCTTGGAGTGAATGAGG 1201

|||

Qy 1201 GAATGATGATGAGAGAGCGCTGAGAGCGCAGAGTGAATGCAACCACTTCTGATGCTG 1260

Db 1202 GAATGATGATGAGAGAGCGCTGAGAGCGCAGAGTGAATGCAACCACTTCTGATGCTG 1261

|||

Qy 1261 AGCGGTGCTGAGAGCGCCCGCCAGCTGGAAGCCCAAGATGCGACCGCCAGAGAG 1320

Db 1262 AGCGGTGCTGAGAGCGCCCGCCAGCTGGAAGCCCAAGATGCGACCGCCAGAGAG 1321

|||

Qy 1321 CCTTTAGACCCCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380

Db 1322 CCTTTAGACCCCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1381

|||

Qy 1381 CCTTGAAGGCTGAGAGGCTTCTCAACACTCTCTGCGGGTGGAGAGAGTCTCTGAGAGA 1440

Db 1382 CCTTGAAGGCTGAGAGGCTTCTCAACACTCTCTGCGGGTGGAGAGAGTCTCTGAGAGA 1441

|||

Qy 1441 AGGTGAGAGTGGAGAGCTATGCTTGGGGGAGAGCGCAGATGAATGCTCTGCTGAG 1500

Db 1442 AGGTGAGAGTGGAGAGCTATGCTTGGGGGAGAGCGCAGATGAATGCTCTGCTGAG 1501

|||

Qy 1501 GAGCCAGCTGCTGGGCGCCAGCTGCTATGTAAGGCTTCTGCTGCTGCTGCTGCTGCT 1560

Db 1502 GAGCCAGCTGCTGGGCGCCAGCTGCTATGTAAGGCTTCTGCTGCTGCTGCTGCTGCT 1561

|||

Qy 1561 CACCCAGCAACCAATCCAAATGAATGATGATCCCA 1599

Db 1562 CACCCAGCAACCAATCCAAATGAATGATGATCCCA 1600

|||

RESULT 2

AK022402 1821 bp mRNA linear PRI 01-AUG-2002

LOCUS Homo sapiens cDNA FLJ12340 fis, clone MAMMA1002268, moderately

DEFINITION similar to Mus musculus spinosinase kinase (SPK1a) mRNA.

ACCESSION AK022402

VERSION AK022402.1 GI:10433790

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens Mammary gland cDNA to mRNA, clone_1lb:MAMMA1

ORGANISM clone:MAMMA1002268.

Homio sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,H., Sudo,H., Wagatsuma,M., Hosokawa,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuo,Y., Niomiya,K. and Iwayanagi,T. NEDO human cDNA sequencing project

TITLE JOURNAL

REFERENCE

2 (bases 1 to 1821)

Isogai,T. and Otsuki,T.

TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory: 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp) Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'-3' end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source Location/Qualifiers
1. 1821
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1002268"
/tissue_type="Mammary gland"
/clone_11b="MAMMA1"
/note="Cloning vector: pME18SFL3"
393. 1547
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14028.1"
/db_xref="GI:10433791"
/translation="MDPAGGPRGVLPKPCRVLYLNPBGKALQIFRSHVPLAE
AEISFTMLTERNHARELVRSEELRMDALVMSGDGLMEHVNGLMRPMETAIO
KPICLPAGSGNALAASLNHYAGEOVNEDLTTCILICRILSPMLLSHTASG
LRPSVISLAEAFIDVLESKRYRLSEMRTLTSTPLRLALRYRRLATLPGRV
GSKTPASPVVVOGFPVDHLVLEPVSHTVDEDEVLVLLHSLGSEMPAP
MGRCAVMHLEFYRAGVSRAALLRLFLAMEGRHMECPPLVLPVPAVFLBEPKDG
KGFVAVDELMSVAVOGVHPNFMVNSGVPEPPSWKPOOMPPEEPL"
BASE COUNT 301 a 569 c 614 g 337 t
ORIGIN
Query Match .97.6%; Score 1560.8; DB 9; Length 1821;
Best Local Similarity 98.6%; Pred. No. 2.5e-265;
Matches 1574; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 3 CCCCCGGCTCTTAAAGCCAGGCTCCGGGCGGGAAGGCCACAGCCGCGCTG 62
DB 221 CCCCCGGCTCTTAAAGCCAGGCTCCGGGCGGGAAGGCCACAGCCGCGCTG 280
QY 63 CGAGCCGCGCTGGGCGAGCAGCATAGAGCTGAAGGAGGAGCGCGCCAGGAG 122
DB 281 CGAGCCGCGCTGGGCGAGCAGCATAGAGCTGAAGGAGGAGCGCGCCAGGAG 340
QY 123 CGCCCCCAGCGGCGAGGAGCCCTTGCGAGCGGGAAGCGCGGCTGAGATTGATCC 182
DB 341 CGCCCCCAGCGGCGAGGAGCCCTTGCGAGCGGGAAGCGCGGCTGAGATTGATCC 400
QY 183 AGCGGGGCGCGCGCGGCGCTGCTCCGGGCGCGCTGCGGCTGCTGCTGAACC 242
DB 401 AGCGGGGCGCGCGCGGCGCTGCTCCGGGCGCGCTGCGGCTGCTGCTGAACC 460
QY 243 GCGGCGGCGCAAGGGAAGGCTTGAGCTCTTCGAGTCACTGTCAGACCCCTTTGGC 302
DB 461 GCGGCGGCGCAAGGGAAGGCTTGAGCTCTTCGAGTCACTGTCAGACCCCTTTGGC 520
QY 303 TGAAGCTGAATCTCCTTACGCTGATGCTCACTGAGCGGGAACACAGCGGGAGCT 362
DB 521 TGAAGCTGAATCTCCTTACGCTGATGCTCACTGAGCGGGAACACAGCGGGAGCT 580
QY 363 GGTGGGCTGGAAGAGCTGGGCGGCTGGGAGCCTGCGTGGTCACTGAGAGAGGCT 422
DB 581 GGTGGGCTGGAAGAGCTGGGCGGCTGGGAGCCTGCGTGGTCACTGAGAGAGGCT 640
QY 423 GATGACAGAGTGTGAACGGGCTCATGAGAGCGGCTGACTGAGAGACCGCATCAGAA 482
DB 641 GATGACAGAGTGTGAACGGGCTCATGAGAGCGGCTGACTGAGAGACCGCATCAGAA 700
QY 483 GCCCCTGTGTAGCTCCAGAGGCTCTGGCAACGGCTGGCAGCTTCTTGAACCATTA 542
DB 701 GCCCCTGTGTAGCTCCAGAGGCTCTGGCAACGGCTGGCAGCTTCTTGAACCATTA 760

QY 543 TCGTGCTATGAGCAGAGTGCACCATATGAAGACCTCTGACACATGCTATGCTGTG 602
DB 761 TCGTGCTATGAGCAGAGTGCACCATATGAAGACCTCTGACACATGCTATGCTGTG 820
QY 603 CCGCCCGGCTGTACACCATATGAAGACCTCTGCTGTGACACAGGCTTGGGCGTGC 662
DB 821 CCGCCCGGCTGTACACCATATGAAGACCTCTGCTGTGACACAGGCTTGGGCGTGC 880
QY 663 GTTCTGTGTGTACAGCTTGGGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGT 722
DB 881 GTTCTGT 940
QY 723 GTATGGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCCGCGCTTGGAGCCCT 782
DB 941 GTATGGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCCGCGCTTGGAGCCCT 1000
QY 783 GCGCACTTACCGCGCGCGAGCTGCTACCTCCCTGTAGAGAGTGGTTTGAAGACACC 842
DB 1001 GCGCACTTACCGCGCGCGAGCTGCTACCTCCCTGTAGAGAGTGGTTTGAAGACACC 1060
QY 843 TCGTTCGCCCGT 902
DB 1061 TCGTTCGCCCGT 1120
QY 903 GCGAGTGCCTTCTCACTGTGAGGCTGCGCCAGCAGAGACTTGTGTGTGTGTGTGT 962
DB 1121 GCGAGTGCCTTCTCACTGTGAGGCTGCGCCAGCAGAGACTTGTGTGTGTGTGTGT 1180
QY 963 GCTGCACCTGCACTGTGGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
DB 1181 GCTGCACCTGCACTGTGGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1240
QY 1023 CGTATGATGATGT 1082
DB 1241 CGTATGATGATGT 1300
QY 1083 CCGTGCAGAGAGAGGAGGAGGAGCATATGAGATGAGATGAGATGAGATGAGAT 1142
DB 1301 CCGTGCAGAGAGAGGAGGAGGAGCATATGAGATGAGATGAGATGAGATGAGAT 1360
QY 1143 CGTGTGCGCTTCCGCTTGTGAGCCCAAGAGATGAGATGAGATGAGATGAGATGAG 1202
DB 1361 CGTGTGCGCTTCCGCTTGTGAGCCCAAGAGATGAGATGAGATGAGATGAGATGAG 1420
QY 1203 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1262
DB 1421 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1480
QY 1263 CGGTTGCGTGTGAGGCGCGCGCGCGAGTGTGAGAGGCGCGAGATGAGAGGCGCG 1322
DB 1481 CGGTTGCGTGTGAGGCGCGCGCGCGAGTGTGAGAGGCGCGAGATGAGAGGCGCG 1540
QY 1323 CTATGAGCCCTGTGGGCGCGCTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGT 1382
DB 1541 CTATGAGCCCTGTGGGCGCGCTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGT 1600
QY 1383 CTAGGCTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1442
DB 1601 CTAGGCTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1660
QY 1443 GGTGAGAGTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1502
DB 1661 GGTGAGAGTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1720
QY 1503 GCCCAGCTGTGTGGGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1562
DB 1721 GCCCAGCTGTGTGGGCGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1780
QY 1563 CCGCAGCAACCAATCAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1598
DB 1781 CCGCAGCAACCAATCAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1816

RESULT 3
AK023393
LOCUS
DEFINITION Homo sapiens cDNA FLJ1331 fls, clone OVARC1001809, moderately
similar to Mus musculus sphingosine kinase (SPHK1a) mRNA.
ACCESSION AK023393
VERSION AK023393.1 GI:10435311
KEYWORDS oligo cloning; fls (full insert sequence).
SOURCE Homo sapiens ovary, tumor tissue cDNA to mRNA, clone_1lb:OVARC1
clone:OVARC1001809.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shitatori,A., Sudo,H.,
Magatsuma,M., Takahashi,M., Chiba,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Matanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Salto,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahari,K., Masuno,Y., Minomiyu,K. and Iwayanagi,T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1772)
Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
1..1772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="OVARC1001809"
/tissue_type="ovary, tumor tissue"
/clone_1lb="OVARC1"
/note="cloning vector: pME18FLJ3"
350..1504
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14558.1"
/db_xref="GI:10435312"
/translation="MDPAGGPRGVLPRLPRLVLLNPRGKSKALQLFRSHVPLLA
AEISFTIMTERNHARELVRSEELGRMDALVMSGDGLMHEVNLMPERPMETAIQ
KPLCSLPAGSNALASLNHYAGYEDVNTEDLLTCTLLICRLISPMNLISLTHASG
ILPLSVLSLAKFPIADVDLESKRYRLGEMRTLGLFRLALRTYRRLASLSEMPRA
GSKTPASPVVVOGPDVDAHLVLEBPVSHWTVVDEDFVLVLLIHLSHISEMRAAP
MORCAAGVHLIFVRAGVSRAMLRLFLAMEGRMEYECPLVTVPAVAFLEPKDG
KGVFAVDGLIMVSEAVOGVHPNIFMNVSGVPEPPSMKPPQMPPEBPL"
BASE COUNT 285 a 566 c 585 g 336 t
ORIGIN
Query Match 97.4%; Score 1558.2; DB 9; Length 1772;
Best Local Similarity 98.6%; Pred. No. 7.2e-265;
Matches 1572; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Db 298 CGCCCCACAGCGCCAGGAGACCCCTGGAGAGGGAGCCGGGCTGAGTTATGATCC 357
Oy 183 AGCGGGGGCCCCCGGGGGCTGCTCCCGGGCCCTCCCGCTGCTGCTCTCAACCC 242
Db 358 AGCGGGGGGGCCCCCGGGGGCTGCTCCCGGGCCCTCCCGCTGCTGCTCTCAACCC 417
Oy 243 GCGGGGGGGAGGGGAGGAGGCTTGGAGCTTCCGGAATCACTGACAGCCCTTTTGGC 302
Db 418 GCGGGGGGGAGGAGGAGGCTTGGAGCTTCCGGAATCACTGACAGCCCTTTTGGC 477
Oy 303 TGAAGCTGAATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
Db 478 TGAGGCTGAATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
Oy 363 GCGGGGGGGAGGAGGAGGCTGCGGGGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
Db 538 GGTGGGCTCGAGAGAGCTGGGGCGGTGGAGAGCTTGTGGTATCTGTGGAGAGCGGCT 597
Oy 423 GATGACAGAGTGTGAACAGGGCTCATGAGAGCGGCTGACTGGAGACCGCATCAGAA 482
Db 598 GATGACAGAGTGTGAACAGGGCTCATGAGAGCGGCTGACTGGAGACCGCATCAGAA 657
Oy 483 GCCCTGTGAGCTTCCACAGAGCTTGGCAACGGCTGGCAAGCTTCTTGAACATTA 542
Db 658 GCCCTGTGAGCTTCCACAGAGCTTGGCAACGGCTGGCAAGCTTCTTGAACATTA 717
Oy 543 TGCTGGGTATGAGAGAGTGCACCAATGAAGAGCTCTGACCAATGACAGCTATGCTGTG 602
Db 718 TGCTGGGTATGAGAGAGTGCACCAATGAAGAGCTCTGACCAATGACAGCTATGCTGTG 777
Oy 603 CCGCGCGGTGCTGTACCCATGAACTGTGTCTTTCACAGAGCTTCCGGGCTGCGCTC 662
Db 778 CCGCGCGGTGCTGTACCCATGAACTGTGTCTTTCACAGAGCTTCCGGGCTGCGCTC 837
Oy 663 GTTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
Db 838 GTTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
Oy 723 GTATCGGCTGTGAGGAGAGTGCCTTCACTGAGGAGCTTCCGCTGTGAGGCTT 782
Db 898 GTATCGGCTGTGAGGAGAGTGCCTTCACTGAGGAGCTTCCGCTGTGAGGCTT 957
Oy 783 GCGGACCTACCGCGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
Db 958 GCGGACCTACCGCGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
Oy 843 TGCTTCCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
Db 1018 TGCTTCCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
Oy 903 GAGGCTGCTTCTCACTGCGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
Db 1078 GCGAGTGGCTTCTCACTGCGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
Oy 963 GCTGACCTGCGAGCTGCGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
Db 1138 GCTGACCTGCGAGCTGCGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
Oy 1023 GCTGATGATCTGCTTCACTGCGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
Db 1198 GCTGATGATCTGCTTCACTGCGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
Oy 1083 CTTGCGCATGAGAGAGGCGAGCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1142
Db 1258 CTTGCGCATGAGAGAGGCGAGCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1317
Oy 1143 CTTGCGCATGAGAGAGGCGAGCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1202
Db 1318 CTTGCGCATGAGAGAGGCGAGCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1377
Oy 1203 ATTGATGATGAGAGAGGCGAGCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAG 1262

Db 1378 ATTGATGGTATGACGAGCGCTGACAGGCGCCAGGTGACCCAACTACTTCTGGATGTGAC 1437

QY 1263 CGGTGCGTGGAGCCCGCCAGCTGGAAGCCCGCAGAGATGCCACCGCCAGAGAGACC 1322

Db 1438 CGGTGCGTGGAGCCCGCCAGCTGGAAGCCCGCAGAGATGCCACCGCCAGAGAGACC 1497

QY 1323 CTTATGACCCCGTGGCGCGCTGCTGCTTATGCTACTTGTGAGAGACCTTCCTCTCC 1382

Db 1498 CTTATGACCCCGTGGCGCGCTGCTGCTTATGCTACTTGTGAGAGACCTTCCTCTCC 1557

QY 1383 CTAGGCGTGCAGAGCGCTGCTGCACAGCTCTCTGCGGTGAGAGACCTCTCTGAGAG 1442

Db 1558 CTAGGCGTGCAGAGCGCTGCTGCACAGCTCTCTGCGGTGAGAGACCTCTCTGAGAG 1617

QY 1443 GGTGAGAGGTGAGAGCTATGCTTGGGGGAGAGAGAGATGAGAGCTGCTGAGAGAG 1502

Db 1618 GGTGAGAGGTGAGAGCTATGCTTGGGGGAGAGAGAGATGAGAGCTGCTGAGAGAG 1677

QY 1503 GCCCAGCTGCTGGCGCCAGCTGCTATGTAAGGCTTCTAGTTGTTCTGAGAGCCCA 1562

Db 1678 GCCCAGCTGCTGGCGCCAGCTGCTATGTAAGGCTTCTAGTTGTTCTGAGAGCCCA 1737

QY 1563 CCCACGAAACCAATCCAAATGAATGACATTCC 1597

Db 1738 CCCACGAAACCAATCCAAATGAATGACATTCC 1772

RESULT 4

LOCUS BC008040 1824 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, clone MGC:15041 IMAGE:3831657, mRNA, complete cds.

ACCESSION BC008040

VERSION BC008040.1 GI:14165485

KEYWORDS MGC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1824)

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

TITLE Strausberg, R.

JOURNAL Direct Submission

REMARK Submitted (21-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

EMAIL: cgapbs-remail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadanesystemsbiology.org

Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan

FEATURES

source 1. 1824

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:15041 IMAGE:3831657"

/tissue_type="Skin, melanotic melanoma."

/clone_id="NIH-MGC_20"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

339..1493

/codon_start=1

CDS

/product="Unknown (protein for MGC:15041)"

/protein_id="AA08040.1"

/db_xref="GI:14165486"

/translation="MDPAGPRGVLPKRCVLYLNPBGKGLQLFRSVQPLAE
AEISFTLTERRNHARELVSBEIGRDALVYMSGDLMEHVNGLMERDMEAIQ
KPLCSLPAGSGNALAASLNHAYGEQVYNEDELNCTLLCLRLLSPNLLSLTASG
LRFSVLSLAFGLADYSEKRYRLGEMFTLGTFTLLAALRYRGLRSLYLVGRV
GSKTPSPVVOOGPYDAHLVPLEPSPVSHMTVDDEFVLVLLHSLGSEMAP
MGRCAQYMHLPYRAGYSRMLRLPLAMERGRMEVECYLYVYVVARLRLEPKDS
KGVRADELAVSEAVOVHNFHMFVSGVEPPSKPMQMPPEEP."

BASE COUNT 354 a 551 c 592 g 327 t

ORIGIN

Query Match 97.4% Score 1557; DB 9; Length 1824;

Best Local Similarity 98.4% Pred No. 1.2e-264;

Matches 1572; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

3 CCCCGGGCTCTATAGCCAGCGCTCCGGGGGGAAGCGAGCCACAGCGGCGCTG 62

167 CGCGGTGCTCTACAGCAGCGCTCCGGGGGGAAGCGAGCCACAGCGGCGCTG 226

63 CGAGCCCGCTTGGGACAGCAGATAGAGAGTGAAGCGAGAGCGGCGGCGAG 122

227 CGAGCCCGCTTGGGACAGCAGATAGAGAGTGAAGCGAGAGCGGCGGCGAG 286

123 CGCCCGCAGCGCGCAGGAGCGCCCTGGGAGCGGAGCGGCGGTGAGATTGATCC 182

287 CGCCCGCAGCGCGCAGGAGCGCCCTGGGAGCGGAGCGGCGGTGAGATTGATCC 346

183 AGCGGGCGGCGCGCGGCGGCTGCTCCGCGGCGCTGCGCGGTGCTGCTGTAACCC 242

347 AGCGGGCGGCGCGCGGCGGCTGCTCCGCGGCGCTGCGCGGTGCTGCTGTAACCC 406

243 GCGCGCGGCGCAAGGCGAGGCTTGCAGCTCTCCGAGTCACTGTCAGCCCTTTGGC 302

407 GCGCGCGGCGCAAGGCGAGGCTTGCAGCTCTCCGAGTCACTGTCAGCCCTTTGGC 466

303 TGAGGCTGAATCTCTTACAGCTGATGCTGCTAGCGGCGGAGCAAGCGGCGGAGCT 362

467 TGAGGCTGAATCTCTTACAGCTGATGCTGCTAGCGGCGGAGCAAGCGGCGGAGCT 526

363 GGTGGGCTGCGAGAGAGTGGGCGGCTGGGAGCGCTGTGTGATGTGTGAGAGGAGCT 422

527 GGTGGGCTGCGAGAGAGTGGGCGGCTGGGAGCGCTGTGTGATGTGTGAGAGGAGCT 586

423 GATGACAGAGTGTGGAAGGCGCTCATGAGAGGCGCTGACTGAGAGAGCCCATCCAGA 482

587 GATGACAGAGTGTGGAAGGCGCTCATGAGAGGCGCTGACTGAGAGAGCCCATCCAGA 646

483 GCGCCCTGTATACCTCCAGAGGCTTGGCAAGCGGCTGGGAGCGCTTCTTGAACATTGA 542

647 GCGCCCTGTATACCTCCAGAGGCTTGGCAAGCGGCTGGGAGCGCTTCTTGAACATTGA 706

543 TGCTGTATAGAGAGGCTTACCAATGAAGAGCTCTGACCAACTGACAGCTATTGCTGTG 602

707 TGCTGTATAGAGAGGCTTACCAATGAAGAGCTCTGACCAACTGACAGCTATTGCTGTG 766

603 CGCGCGGCTGTATACCAATGAAGAGCTCTGACCAACTGACAGCTATTGCTGTG 662

767 CGCGCGGCTGTATACCAATGAAGAGCTCTGACCAACTGACAGCTATTGCTGTG 826

663 GTTCTGTGTGACAGCTTGGGCGCTTATGCTGATGAGAGCTTGAAGAGATGA 722

827 GTTCTGTGTGACAGCTTGGGCGCTTATGCTGATGAGAGCTTGAAGAGATGA 886

723 GTATGCGGCTTGGGAGAGATGCGCTTACCTTGGGAGCTTCTGCTTGGAGAGCT 782

887 GTATGCGGCTTGGGAGAGATGCGCTTACCTTGGGAGCTTCTGCTTGGAGAGCT 946

783 GGGACACTACCGCGGCGAGTGGCTACCTTCTTGAAGAGAGTGGTTCAACAGCC 842

947 GCGACACTACCGCGGCGAGTGGCTACCTTCTTGAAGAGAGTGGTTCAACAGCC 1006

QY 843 TGGTTCCTCCCGTTGGTTCAGAGAGGCGCCGGTAGATGACACCTGGTGGCCACTGGAGGA 902
 DB 1007 TGCTTCCCGCTGGTTGGTTCAGAGAGGCGCCGGTAGATGACACCTGGTGGCCACTGGAGGA 1066
 QY 903 GCAGTGCCTTTCCTCAGTGGAGAGTGTGCCCCAGAGAGACTTGTGTCTAGTCTGGGACT 962
 DB 1067 GCCAGTGCCTTTCCTCAGTGGAGAGTGTGCCCCAGAGAGACTTGTGTCTAGTCTGGGACT 1126
 QY 963 GCTGCATCTGCAGCTGGCCAGTGTGATGTTGCTGCGACCCATGGGCGGCTGTGAGTGG 1022
 DB 1127 GCTGCATCTGCAGCTGGCCAGTGTGATGTTGCTGCGACCCATGGGCGGCTGTGAGTGG 1186
 QY 1023 CGTATGCATCTGTTTACAGTGGCGGGGAGTGTGCTGCGACCTGCTGCGCTCTT 1082
 DB 1187 CGTATGCATCTGTTTACAGTGGCGGGGAGTGTGCTGCGACCTGCTGCGCTCTT 1246
 QY 1083 CCTGCGCATGAGAGAGGCGAGCATATGAGTATGAATGCCCTACTTGTGTATATGTTGCC 1142
 DB 1247 CCTGCGCATGAGAGAGGCGAGCATATGAGTATGAATGCCCTACTTGTGTATATGTTGCC 1306
 QY 1143 CGTGTGCGCTTCCGCTTGGAGGCGGAGGATGGGAAAGGTGTGTTGCAAGTGGGGA 1202
 DB 1307 CGTGTGCGCTTCCGCTTGGAGGCGGAGGATGGGAAAGGTGTGTTGCAAGTGGGGA 1366
 QY 1203 ATTGATGTTAGCAGAGGCGCTGTCAGAGGCGCAGGTGCACCCCAACTACTTGTGATGTCAG 1262
 DB 1367 ATTGATGTTAGCAGAGGCGCTGTCAGAGGCGCAGGTGCACCCCAACTACTTGTGATGTCAG 1426
 QY 1263 CGGTTGCGTGGAGGCGGCGCCAGCTGGAAGCCCCAGACATATGCGACCGCCAGAGAGGCC 1322
 DB 1427 CGGTTGCGTGGAGGCGGCGCCAGCTGGAAGCCCCAGACATATGCGACCGCCAGAGAGGCC 1486
 QY 1323 CTTATGACCCCTGGGCGCGCTGTGCTTACTGTTACTGTCAGAGACCCCTGCTCTCTCC 1382
 DB 1487 CTTATGACCCCTGGGCGCGCTGTGCTTACTGTTACTGTCAGAGACCCCTGCTCTCTCC 1546
 QY 1383 CTAGGCGTGCAGAGGCGCTGTCACAGCTCTCTGTGGGGGTGAGAGAGACTCTGTGAGAG 1442
 DB 1547 CTAGGCGTGCAGAGGCGCTGTCACAGCTCTCTGTGGGGGTGAGAGAGACTCTGTGAGAG 1606
 QY 1443 GGTGAGAGGTGGAGGCTATGCTTTGGGGGAGACAGGCGGAGATGAATGATCTGTGGTCA 1502
 DB 1607 GGTGAGAGGTGGAGGCTATGCTTTGGGGGAGACAGGCGGAGATGAATGATCTGTGGTCA 1666
 QY 1503 GCCCGAGTGGTGGGCGCCAGCTGCTATGTAAGGCTCTGATTTGTTCTAGAGACCCCA 1562
 DB 1667 GCCCGAGTGGTGGGCGCCAGCTGCTATGTAAGGCTCTGATTTGTTCTAGAGACCCCA 1726
 QY 1563 CCCGACGAGCAAAATCAAAATTAAGTGAATTCCTCCAA 1599
 DB 1727 CCCGACGAGCAAAATCAAAATTAAGTGAATTCCTCCAA 1763
 * RESULT 5
 AF238083 1783 bp mRNA linear PRI 01-JUN-2000
 LOCUS AF238083 Homo sapiens sphingosine kinase-1 mRNA, complete cds.
 DEFINITION AF238083
 ACCESSION AF238083.1 GI:8132867
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1783)
 AUTHORS Nava,V.E., Lacana,E., Poulton,S., Liu,H., Sugitara,M., Kono,K.,
 Milstien,S., Kohama,T. and Spiegel,S.
 TITLE Functional characterization of human sphingosine kinase-1
 JOURNAL FEBS Lett. 473 (1), 81-84 (2000)
 MEDLINE 20263733
 PUBMED 10802064
 REFERENCE 2 (bases 1 to 1783)
 AUTHORS Nava,V.E., Lacana,E., Poulton,S., Liu,H., Sugitara,M., Kono,K.,

Milstien,S., Kohama,T. and Spiegel,S.
 Direct Submission
 Submitted (23-FEB-2000) Biochemistry, Georgetown University, 3900
 Reservoir RD NW BSB Rm 357, Washington, DC 20007, USA
 FEATURES
 source
 1. 1783
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q25.2"
 337, .1491
 /codon_start=1
 /product="sphingosine kinase-1"
 /protein_id="AF238083.1"
 /db_xref="GI:8132868"
 /translation="MDPAGGPRGVLPRPCRVLYLNPGRGKGLOLFPRSHVPLAE
 AETISFLMTERNNHRELYRSELEGRNALVYMSDGLMHEVUNGIMRPMETRAIO
 KPLCSLPAGSGNGLASLNHYAGTEQVITNEDLLNCTLLCRLLSPMLLSLHTRSG
 LRLESVLSLAMEFIADVDESEKRYRLGEMRTLLGLFLALRLYRGLAYLPGRV
 GSKTPASPVVVOGPDVADHLYPLEEPVPSHTWMPDEDEVLILALSHLSGEMFAP
 MGRCAAGVHILFYRAGVSRAMILREFLAMEGRHMEYCPYLIVYVPAVFLERPDG
 KGVFAYDGLMVEAVOGVHPHYFMVYSGVPEPSPMRPQMPPEEPL"
 CDS
 314 a 546 c 591 g 332 t
 BASE COUNT
 ORIGIN
 Query Match 96.9%; Score 1549.6; DB 9; Length 1783;
 Best local similarity 98.2%; Pred. No. 2,4e-263;
 Matches 1567; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
 QY 3 CCCCGGGGCTCCTATACGACGCGCTCGGGGGGGAAGGAGACCCGACAGCGGGCGCTG 62
 DB 165 CACCGGTGCTCTACACGCGCTCGGGGGGGAAGGAGACCCGACAGCGGGCGCTG 224
 QY 63 CGAGCGCGCTGGGCGAGCACCGATAGAGAGTGAAGGAGAGCCGCGCACGCGGAG 122
 DB 225 CGAGCGCGCTGGGCGAGCACCGATAGAGAGTGAAGGAGAGCCGCGCACGCGGAG 284
 QY 123 CGCGCCACAGCGCGACGCGCTGCGAGCGGAGACCGCGGCTGAGGTTATGATCC 182
 DB 285 CGCGCCACAGCGCGACGCGCTGCGAGCGGAGACCGCGGCTGAGGTTATGATCC 344
 QY 183 AGCGGGGCGCGCGGGGCGTGCCTCGCGGCGCTCGCGGCTGCTGCTGAACCC 242
 DB 345 AGCGGGGCGCGCGGGGCGTGCCTCGCGGCGCTCGCGGCTGCTGCTGAACCC 404
 QY 243 GCGCGGCGCAAGGAGCGCTTGCAGCTCTTCCGAGTCACTGAGAGCCCTTTTGGC 302
 DB 405 GCGCGGCGCAAGGAGCGCTTGCAGCTCTTCCGAGTCACTGAGAGCCCTTTTGGC 464
 QY 303 TGAGGCTGAATCTCTTCAAGCTGATGCTACAGCGCGGAGACAGCGCGGAGCT 362
 DB 465 TGAGGCTGAATCTCTTCAAGCTGATGCTACAGCGCGGAGACAGCGCGGAGCT 524
 QY 363 GGTGCGGTGGAGAGCTGGGCGCGTGGGAGCGCTGTGTCATGCTGTGAGAGGCT 422
 DB 525 GGTGCGGTGGAGAGCTGGGCGCGTGGGAGCGCTGTGTCATGCTGTGAGAGGCT 584
 QY 423 GATCAGAGAGTGTGAAGCGGCTCATGAGAGCGGCGCTGACTGGAGAGCGGCATCCAGAA 482
 DB 585 GATCAGAGAGTGTGAAGCGGCTCATGAGAGCGGCGCTGACTGGAGAGCGGCATCCAGAA 644
 QY 483 GCCCGTGTAGCTCCAGAGGCTGTGGCAACGCGCTGGCAGCTTCTTGAACCATTA 542
 DB 645 GCCCGTGTAGCTCCAGAGGCTGTGGCAACGCGCTGGCAGCTTCTTGAACCATTA 704
 QY 543 TGTGTGTATGAGAGGTGCACCAATGAAGACCTCTGACCAACGACGCTATTTGCTGTG 602
 DB 705 TGTGTGTATGAGAGGTGCACCAATGAAGACCTCTGACCAACGACGCTATTTGCTGTG 764
 QY 603 CCGCGCGGTGCTGCACCATGAAGCTGTGCTCTGCACAGCGCTTGGGGGTGCGCTC 662
 DB 765 CCGCGCGGTGCTGCACCATGAAGCTGTGCTCTGCACAGCGCTTGGGGGTGCGCTC 824

663 GTTCTGTGTGCTGAGCTGGCTGGGGCTTATGCTGATGTGAGCACTGAGAGTATTA 722
|||||
Db 825 CTTCCTGTGTCTGACAGCTGGCTGGGGCTTATGCTGATGTGAGCACTGAGAGTATTA 884
723 GATATGGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCTGCTGTGGACGCTT 782
|||||
Db 885 GATATGGGCTGTGGGGAGATGCGCTTCACTGTGGGACCTTCTGCTGTGGACGCTT 944
783 GGGCACTTACCGCGCGGCGGCTGCTACCTCCCTGTAGAGAAAGTGGTTTCAAGACAC 842
|||||
Db 945 GGGCACTTACCGCGCGGCGGCTGCTACCTCCCTGTAGAGAAAGTGGTTTCAAGACAC 1004
843 TGTCTTCCCGCTGTGTGTGCTGACAGAGCGCGGTAGATGACACCTGTGCACTGGAGA 902
|||||
Db 1005 TGTCTTCCCGCTGTGTGTGCTGACAGAGCGCGGTAGATGACACCTGTGCACTGGAGA 1064
903 GCGAGTGGCTTGTCTCACTGGAGGTGTGCGCCGACGAGACTTGTGTGCTAGTCTGGACT 962
|||||
Db 1065 GCGAGTGGCTTGTCTCACTGGAGGTGTGCGCCGACGAGACTTGTGTGCTAGTCTGGACT 1124
963 GGTGACCTGACACTGTGGCAGTGAATGTTTGTCTGACACCATGGCGCTGTGCACTGG 1022
|||||
Db 1125 GGTGACCTGACACTGTGGCAGTGAATGTTTGTCTGACACCATGGCGCTGTGCACTGG 1184
1023 CGTCAATGATCTGTCTTACGTGGGCGGAGAGTGTCTGTCATGCTGTGGCTT 1082
|||||
Db 1185 CGTCAATGATCTGTCTTACGTGGGCGGAGAGTGTCTGTCATGCTGTGGCTT 1244
1083 CCTGGCCATGAGAAAGGAGGAGCATATGAGATATGAAATGCCCTTACTTGTATATGTGC 1142
|||||
Db 1245 CCTGGCCATGAGAAAGGAGGAGCATATGAGATATGAAATGCCCTTACTTGTATATGTGC 1304
1143 CGTGTGCGCTTCCCTGTGTGAGCCCAAGATGGGAAAGTGTGTGTGCACTGGATGGCA 1202
|||||
Db 1305 CGTGTGCGCTTCCCTGTGTGAGCCCAAGATGGGAAAGTGTGTGTGCACTGGATGGCA 1364
1203 ATTGATGTGTAGCAGAGCGGTGTCAGAGGCGCAGGTGACCCCAACTACTTGTGATGTGAG 1262
|||||
Db 1365 ATTGATGTGTAGCAGAGCGGTGTCAGAGGCGCAGGTGACCCCAACTACTTGTGATGTGAG 1424
1263 GGTGTGCGTGTGAGAGCCCGCCCAAGTGTGAAGCCCAAGACAGATGCCACCGCCAGAGAGCC 1322
|||||
Db 1425 TGTGTGCGTGTGAGAGCCCGCCCAAGTGTGAAGCCCAAGACAGATGCCACCGCCAGAGAGCC 1484
1323 CTATGAGCCCTGTGGGCGGCGGTGTGCTTATGTCTACTTGTGAGAGCCCTTCTCTCTTC 1382
|||||
Db 1485 CTATGAGCCCTGTGGGCGGCGGTGTGCTTATGTCTACTTGTGAGAGCCCTTCTCTCTTC 1544
1383 CTAGGCTGTGAGAGCGCTGTCAAGCTCTGTGTGGGCTGAGAGACTCTCTGTGAGAG 1442
|||||
Db 1545 CTAGGCTGTGAGAGCGCTGTCAAGCTCTGTGTGGGCTGAGAGACTCTCTGTGAGAG 1604
1443 GGTGAGAGGTGAGAGCTATGCTTGTGGGGGACAGGCCAAGTGAAGTCTTGGGTGAGA 1502
|||||
Db 1605 GGTGAGAGGTGAGAGCTATGCTTGTGGGGGACAGGCCAAGTGAAGTCTTGGGTGAGA 1664
1503 GGGCAGCTGGCTGGGCGGCGGCTGATGTAGAGCCCTTCTAGTTTGTGTGAGACCCCA 1562
|||||
Db 1665 GGGCAGCTGGCTGGGCGGCGGCTGATGTAGAGCCCTTCTAGTTTGTGTGAGACCCCA 1724
1563 CCCCAAGAACCAATCAATTAAGTGAATTCCTCA 1598
|||||
Db 1725 CCCCAAGAACCAATCAATTAAGTGAATTCCTCA 1760

RESULT 6
AX127641 1719 bp DNA 1linear PAT 15-MAY-2001
LOCUS AX127641
DEFINITION Sequence 1 from Patent WO0131029.
ACCESSION AX127641
VERSION AX127641.1 GI:14134307
KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 1719)
AUTHORS Allen, J., Gosink, M., Melendez, A. J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 1 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source location/Qualifiers
1..1719
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 305 a 529 c 556 g 329 t
ORIGIN
Query Match 94.8%; Score 1516.2; DB 6; Length 1719;
Best Local Similarity 97.7%; Pred. No. 1.8e-257;
Matches 1560; Conservative 0; Mismatches 33; Indels 4; Gaps 2;
QY 3 CCCCCGGGCTCCTATATACACAGCGCTCCGGGGCGGGAAGGCCACACAGCCGCGCTG 62
|||||
Db 102 CACCGGTGCTCTTACAGCCAGCGCTCCGGGCGGGAAGGAGGCCACAG--CGGCCCTG 158
63 CGACGCCGCTGTGGGAGCAGCATTAAGAGCTGAAGGAGGAGCGCGCCAGCGAGAG 122
159 CGACGCCGCTGTGGGAGCAGCATTAAGAGCTGAAGGAGGAGCGCGCCAG--GGCAG 217
QY 123 CGCCCCACAGCGCAGGAGGCCCTGTGAGCGGAGGCCGCGGTGAGATTATGATTC 182
|||||
Db 218 CGCCCCACAGCGCAGGAGGCCCTGTGAGCGGAGGCCGCGGTGAGATTATGATTC 277
183 AGCGGGCGGCGCGCGGCGGTGCTCCGGGCGCGCGCGGTGAGCGGTGAGCGGTCAACCC 242
278 AGCGGGCGGCGCGCGGCGGTGCTCCGGGCGCGCGGTGAGCGGTGAGCGGTGAGCGGT 337
QY 243 GCGGCGGCGCAAGGCGCAGGAGCGCTTGTGAGCTTCCGAGTCACTGACAGCCCTTTGGC 302
|||||
Db 338 GCGGCGGCGCAAGGCGCAGGAGCGCTTGTGAGCTTCCGAGTCACTGACAGCCCTTTGGC 397
303 TGAAGCTGAATCTCCTTACAGCTGATGCTCACTAGAGCGGCGGAGAGCT 362
|||||
Db 398 TGAAGCTGAATCTCCTTACAGCTGATGCTCACTAGAGCGGCGGAGAGCT 457
363 GGTGCGGTGAGAGAGCTGGGCGGCGGAGAGCTGTGGGTGATGTGTGAGAGAGGCT 422
458 GGTGCGGTGAGAGAGCTGGGCGGCGGAGAGCTGTGGGTGATGTGTGAGAGAGGCT 517
QY 423 GATGACGAGGTGTGTAAGGAGGCTCATGAGAGCGGCTGACTGGAGACCGCATCAGAA 482
|||||
Db 518 GATGACGAGGTGTGTAAGGAGGCTCATGAGAGCGGCTGACTGGAGACCGCATCAGAA 577
483 GCCCCTGTGTACCTTCCAGAGAGGCTGTGGCAAGCGCTGGCAAGCTTCTTAACCATTA 542
|||||
Db 578 GCCCCTGTGTACCTTCCAGAGAGGCTGTGGCAAGCGCTGGCAAGCTTCTTAACCATTA 637
543 TGTGTGCTATGAGAGGAGTACCAATGAAGACCTCCGACCACTGACGCTATGTGTG 602
|||||
Db 638 TGTGTGCTATGAGAGGAGTACCAATGAAGACCTCCGACCACTGACGCTATGTGTG 697
603 CCGCCGCGGTGTGTACCCATGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 662
698 CCGCCGCGGTGTGTACCCATGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 757
QY 663 GTTCTGT 722
|||||
Db 758 GTTCTGT 817
723 GTATGGCGGTGTGGGGAGATGCGTTCACTGTGGGACCTTCTGCTGTGGACGCTT 782
|||||
Db 818 GTATGGCGGTGTGGGGAGATGCGTTCACTGTGGGACCTTCTGCTGTGGACGCTT 877
783 GCGCACTTACCGGCGGAGCTGCTACCTTCTGTAGAGAGTGGTTTCAAGACAC 842
|||||

Db	878	GCGCACCTACCGCGGCGCAGCTGGGCTACCTCCCTGTAAGAAAGTGGGTTCCAGACACC	937
OY	843	TGCTTCCCCCGTTTGTGTCCAGCAGGGCCGGTGTGATGACACACTGGTGCCTTGAGGA	902
Db	938	TGCCTTCCCCCGTTTGTGTGTCCAGCAGGGCCGGTGTGATGACACACTGGTGCCTTGAGGA	997
OY	903	GCAGGTGCGCTTCTACTGACGAGGGTGGTCCCGACGAGAGACTTGTGCTAGTCCGGGACT	962
Db	998	GCCAGTGCCTTCTACTGACGAGGGTGGTCCCGACGAGAGACTTGTGCTAGTCCGGGACT	1057
OY	963	GCTGCACCTCGCACCCTGGCGCAGTGGAGATGTTTGTGCACCCATGAGGCGCGTGTGCAGCTGG	1022
Db	1058	GCTGCACCTCGCACCCTGGCGCAGTGGAGATGTTTGTGCACCCATGAGGCGCGTGTGCAGCTGG	1117
OY	1023	CGTATGATGATCTGTTTCTACGTCGGGGCGGGAGTGTCTGTGCCATGCTGCTGCCCTCTT	1082
Db	1118	CGTATGATGATCTGTTTCTACGTCGGGGCGGGAGTGTCTGTGCCATGCTGCTGCCCTCTT	1177
OY	1083	CGTGGCCATGAGAGAAGGGCAGCAGTATGAGATATGAATGCCCCACTTGTGATATGTGCC	1142
Db	1178	CCTGGCCATGAGAGAAGGGCAGCAGTATGAGATATGAATGCCCCACTTGTGATATGTGCC	1237
OY	1143	CGTGGTGCCTTCCGCTTGGAGGCCCAAGATGGGAAAGTGTGTTTGACATGTGATGGGGA	1202
Db	1238	CGTGGTGCCTTCCGCTTGGAGGCCCAAGATGGGAAAGTGTGTTTGACATGTGATGGGGA	1297
OY	1203	ATTGATGTTATGACAGCGCCGTGCAGAGGGCCAGGTGCACCCAACTACTCTGTGATGTGAC	1262
Db	1298	ATTGATGTTATGACAGCGCCGTGCAGAGGGCCAGGTGCACCCAACTACTCTGTGATGTGAC	1357
OY	1263	CGGTTTGCCTGGAGCCCCCGCCACGTGTGAGAACCCCAAGATGCCACCGCAGAAAGCC	1322
Db	1358	CGGTTTGCCTGGAGCCCCCGCCACGTGTGAGAACCCCAAGATGCCACCGCAGAAAGCC	1417
OY	1323	CTTATGACCCCTGGGGCCGCGCTGTGCTTACTGTACTTTCAGAGACCCTTCTACTCTCC	1382
Db	1418	CTTATGACCCCTGGGGCCGCGCTGTGCTTACTGTACTTTCAGAGACCCTTCTACTCTCC	1477
OY	1383	CTAGGGCTGCAGGGGCTGTGCACAGCTCTGTGGGGGTGAGAGACTCTCTGGAGAAG	1442
Db	1478	CTAGGGCTGCAGGGGCTGTGCACAGCTCTGTGGGGGTGAGAGACTCTCTGGAGAAG	1537
OY	1443	GGTGGAAGGTGGAGGGCATGCTTTTGGGGGGGACAGGGCCAGAAATGAATCGTGGGTCCAGGA	1502
Db	1538	GGTGGAAGGTGGAGGGCATGCTTTTGGGGGGGACAGGGCCAGAAATGAATCGTGGGTCCAGGA	1597
OY	1503	GCCAGCTGCTGGGGCCAGCTGCTATGTAAAGGCTTCTAGTTTGTCTGAGACCCCA	1562
Db	1598	GCCAGCTGCTGGGGCCAGCTGCTATGTAAAGGCTTCTAGTTTGTCTGAGACCCCA	1657
OY	1553	CCCCACGACCAATCCAAATTAAGTGACATTTCCCA	1599
Db	1658	CCCCACGACCAATCCAAATTAAGTGACATTTCCCA	1694
RESULT 7			
LOCUS	AF266756	1693 bp	mRNA linear PRI 01-JUN-2000
DEFINITION	Homo sapiens sphingosine kinase (SPHK1) mRNA, complete cds.		
ACCESSION	AF266756		
VERSION	AF266756.1	GI:8133099	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1693)		
AUTHORS	Melendez,A.J., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.		
TITLE	Human Sphingosine Kinase, Molecular Cloning, Functional Characterisation and Tissue Distribution		
JOURNAL	Gene (2000) In press		
REFERENCE	2 (bases 1 to 1693)		
AUTHORS	Melendez,A.J., Carlos-Dias,E., Gosink,M., Allen,J.M. and Takacs,L.		

TITLE Direct Submission
 JOURNAL Submitted (11-MAY-2000) Department of Molecular and Cellular Biology, Jouveinal Parke-Davis, 11-13 rue de la Loge, Fresnes 94265, France

FEATURES
 source Location/Qualifiers
 1..1693 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="1/q25.2"
 1..1693 /gene="SPHK1"
 270..1424 /gene="SPHK1"
 /codon_start=1 /product="sphingosine kinase"
 /protein_id="AAF3470.1" /translation="MDPAGGPGRYLPFRPCRYLVLLINPGRGKGAQLQLEFRSHOPLAE
 AEISPTMLTERRHARRELVRSSEEDRNDALVNSGDIMLEVYNGLSMERDPTAIO
 KLPSLIPAGSGNALAASLNHAYGEQVNDLINCILLGRLLSPNNILSLDASG
 LRISVSLIANGFTADYDLSEKXRLRGEMEPFTIGTPIRLALETNRGLAYLVPGVR
 GSKIPASPVYVQGGPVDNHLVPLEPPVSHTYVYPDEFEVLVLLSHLSSEPAAR
 MGRCAAGVMLFVYRAGYSRAMLLRLFLAMKRGHMECEYLVYVPAVRLRLEPKQG
 KGVAIVDELIVSAVQGVHPNFTFMVSGVEPPSPWKPQOMPPEEPL"
 1676..1681 /gene="SPHK1"

polyA_signal
 BASE COUNT 279 a 529 c 556 g 329 t
 ORIGIN

Query Match 94.8%; Score 1515.2; DB 9; Length 1693;
 Best Local Similarity 97.7%; Pred. No. 257;
 Matches 1559; Conservative 0; Mismatches 33; Indels 4; Gaps 2.

QY 3 CCCCCGGGCTCCTATAGCCACGGCTCCGGGGGGGAAGCCGACCCACAGCCGGCCCTG 62
 Db 102 CACCGGTGCTCTTACAGCCACGGCTCCGGGGGAAGGAGACCCACAG---CGCCCTG 158
 QY 63 CGAGCCCGCCTGGGCGACACCGATAGAGCTGAAGGCAGAGCCCGCCACCGGCGAG 122
 Db 159 CGAGCCCGCCTGGGCGAGACCGATAGAGAACTGAAGGCAGAGCCCGCCGAC--GCGAG 217
 QY 123 CGCCCCACAGGCGCAGGAGACCCCTGGCAGCGGGAGCGCGGGGCGAGGTTATGGATCC 182
 Db 218 CGCCCCACAGGCGCAGGAGACCCCTGGCAGCGGGAGCGCGGGGCGAGGTTATGGATCC 277
 QY 183 AGCGGGCGGCCCCCGGGGGCGTGCTCCCGCGGCCCTGCCGCTGTGCTGCTGTGAACC 242
 Db 278 AGCGGGCGGCCCCCGGGGGCGTGCTCCCGCGGCCCTGCCGCTGTGCTGCTGTGAACC 337
 QY 243 GCGCGGGCGGCAAGGCAAGGCTTTCAGACTTTCGGAGTCACTGTGCAAGCCCTTTTGGC 302
 Db 338 GCGCGGGCGGCAAGGCAAGGCTTTCAGACTTTCGGAGTCACTGTGCAAGCCCTTTTGGC 397
 QY 303 TGAGGCTGAATCTCTCTTACGCTATGCTCACTAGAGGGGGAACACAGCGGGAGACT 362
 Db 398 TGAGGCTGAATCTCTCTTACGCTATGCTCACTAGAGGGGGAACACAGCGGGAGACT 457
 QY 363 GGTGCGGTGAGAGAGCTGGGGCGCTGGAGAGCTTGTGTGTCATGTCTGAGAGCGGCT 422
 Db 458 GGTGCGGTGAGAGAGCTGGGGCGCTGGAGAGCTTGTGTGTCATGTCTGAGAGCGGCT 517
 QY 423 GATGACAGAGGGGGAAGCGGCTCATGAGAGGGCTCATGAGGAGACCGGCATTCAGAA 482
 Db 518 GATGACAGAGGGGGAAGCGGCTCATGAGAGGGCTCATGAGGAGACCGGCATTCAGAA 577
 QY 483 GCCCCTGTGTAGCCTCCAGCAGAGCTTGGCAACGCGCTGCGAGCTTCTTGAACATTA 542
 Db 578 GCCCCTGTGTAGCCTCCAGCAGAGCTTGGCAACGCGCTGCGAGCTTCTTGAACATTA 637
 QY 543 TGTGTGGCTATAGCAGAGTCAACAAAGAGACTCTGTACCAACAGTCAAGCTATGTCTGTG 602
 Db 638 TGTGTGGCTATAGCAGAGTCAACAAAGAGACTCTGTACCAACAGTCAAGCTATGTCTGTG 697

/translation="MDPAGRGVRLRLVILNPRGKGAQLQERSHVQPLAEAEI
 SFTLMLERHNRHRELVRSEELGRMDALVYMGSDGMHEVNGLMRDPMEALQKPL
 CSLPAGSGKPLSLNHYAGEVQVNTEDLLCRLRLSPMLLSLHNASGLRL
 PSYLSLAMPFIADYDLSEKRLRGLRGLRLRLRLRLRLRLRLRLRLRLRLRL
 TPASPVYQCGPYDAHLVPLEEPVPSHMYVPEDEFLVILHSLHSDMFAAPGR
 CAAGVHLFTVYRAGVSRAMLRLFLAMEGRHMYECPYLVYVPAFLPEPDKGV
 FAVDELMSEAVQGVHFNTEFMVSGCEVPPSKMFPQMPPEEPL"

BASE COUNT 247 a 422 c 462 g 297 t
 ORIGIN
 Query Match 83.9%; Score 1342.2; DB 9; Length 1428;
 Best Local Similarity 98.3%; Pred. No. 8.4e-227;
 Matches 1356; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 220 CGCGTGTGTGTGTGTGTGAACCGCGCGCGGCAAGGCGCTTGCAGCTTCTCCG 219
 DB 37 CGCGTGTGTGTGTGTGTGAACCGCGCGGCAAGGCGCTTGCAGCTTCTCCG 96
 QY 280 AGTACGTCGACCGCCCTTTGGCTGAGGCTGAATCTCTTCAACGCTGATGCTCAG 339
 DB 97 AGTCAGTCGACCGCCCTTTGGCTGAGGCTGAATCTCTTCAACGCTGATGCTCAG 156
 QY 340 CGCGGGAACACACGCGCGGAGCTGTGGGTGAGGAGCTGGCGCTGAGGAGCTCTG 399
 DB 157 CGCGGGAACACACGCGCGGAGCTGTGGGTGAGGAGCTGGCGCTGAGGAGCTCTG 216
 QY 400 GTGTGATGT 459
 DB 217 GTGTGATGT 276
 QY 460 GACTGGGAGACCGCGCTTCAAGACCGCTGTGTAGCTTCCAGCAGGCTCTGCGACG 519
 DB 277 GACTGGGAGACCGCGCTTCAAGACCGCTGTGTAGCTTCCAGCAGGCTCTGCGACG 336
 QY 520 CTGCGACCTTCTTGAACCATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
 DB 337 CTGCGACCTTCTTGAACCATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 396
 QY 580 ACCAATGACGACCTTGT 639
 DB 397 ACCAATGACGACCTTGT 456
 QY 640 CACACGCGCTTGGGCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 699
 DB 457 CACACGCGCTTGGGCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 516
 QY 700 GATGT 759
 DB 517 GATGT 576
 QY 760 ACCCTCTGCTGT 819
 DB 577 ACCCTCTGCTGT 636
 QY 820 GGAAGGT 879
 DB 637 GGAAGGT 696
 QY 880 GCACACCTGT 939
 DB 697 GCACACCTGT 756
 QY 940 GACTTGT 999
 DB 757 GACTTGT 816
 QY 1000 CCCATGGCGCGCTGT 1059
 DB 817 CCCATGGCGCGCTGT 876
 QY 1060 CGTGCATGT 1119
 DB 877 CGTGCATGT 936

QY 1120 TCCCTCTACTTGT 1179
 DB 937 TCCCTCTACTTGT 996
 QY 1180 GGT 1239
 DB 997 GGT 1056
 QY 1240 CCAAACTACTTGT 1299
 DB 1057 CCAAACTACTTGT 1116
 QY 1300 CAGATGCCACCGCGCAAGAGAGCGCTTATGACCGCTGTGGCGCGCTGTGTGTGT 1359
 DB 1117 CAGATGCCACCGCGCAAGAGAGCGCTTATGACCGCTGTGGCGCGCTGTGTGTGT 1176
 QY 1360 CTTCGAGACCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1419
 DB 1177 CTTCGAGACCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1236
 QY 1420 TGGAGAGACCT 1479
 DB 1237 TGGAGAGACCT 1296
 QY 1480 CAGAATGAAGTCTGT 1539
 DB 1297 CAGAATGAAGTCTGT 1356
 QY 1540 TCTAGTTTGT 1598
 DB 1357 TCTAGTTTGT 1415

RESULT 11
 AB046025 2015_bp mRNA linear PRI 18-JUL-2000
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM
 Macaca fascicularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Macaca.
 1 (sites)
 Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,
 Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
 Isolation of full-length cDNA clones from macaque brain cDNA
 libraries
 Unpublished
 2 (bases 1 to 2015)
 Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
 Direct Submission
 Submitted (14-JUL-2000) Katsuyuki Hashimoto, National Institute of
 Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
 1-chome, Shirojuku-ku, Tokyo 162-8640, Japan
 (E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/
 URL: http://www.nih.go.jp/yoken/genebank/
 Ldb Name: macaque brain cDNA library QcCE
 Lab host: TOP10
 Vector: pME18S-FU3 (ACC. NO. AB00964)
 R. Site1: DraIII (CACTGTGTC)
 R. Site2: DraIII (CACTGTGTC)
 Description: 1st strand cDNA was primed with an oligo(dT) primer
 using specific 5' and 3' primers and amplified by PCR. The PCR
 product was digested with SfiI and size selection was performed to
 exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
 into distinct DraIII sites of pME18S-FU3. XhoI sites just outside

REFERENCE
 JOURNAL
 TITLE
 AUTHORS
 JOURNAL
 TITLE
 JOURNAL
 COMMENT

the Draili sites can be used to isolate the cDNA insert. Libraries were constructed by Sugano et al. (University of Tokyo, Institute of Medical Science). Custom primer used for sequencing (5' end primer [CTTCTGCTCTAAAGCTGG]; 3' end primer [CGACTGCACTCGAGACA]).

FEATURES

source

1. 2015
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="Ocece-16602"
/clone_1id="macaque brain cDNA library Ocece"
/dev_stage="adult"
628..1617
/codon_start=1
/product="unnamed protein product"
/protein_id="BAB01607.1"
/db_xref="GI:9280106"

CDS

BASE COUNT 355 a 639 c 650 g 371 t
ORIGIN
Query Match 82.2%; Score 1314.8; DB 9; Length 2015;
Best Local Similarity 91.4%; Pred. No. 5.4e-222;
Matches 1460; Conservative 0; Mismatches 82; Indels 56; Gaps 4;

OY 3 CCGCGGGGCTCTATAGCCAGCGCTCCGGCGGGGAAAGCCAGCCCGGCGCTG 62
DB 456 CGCGGCTCTCTAGCCAGCGGCTCGGAGGGGAGGCGCCAGCCAGCGGCGCTG 515
OY 63 CGAGCGCGCTCTGGGACACCGATAGAGCTGAAGCAGAGAGCGCGCGCGAG 122
DB 516 CGAGCGCGCTCTGGGACACCGATAGAGCTGAAGCAGAGAGCGCGCGCGAG 575
OY 123 CGCGCGCGCGCGCGCGCGCTGGGAGGAGCGGCGGCGGAGGTTATGATTC 182
DB 576 TGCGCGCGCGCGCGCGCGCGCGCGAGAGGAGCGGCGGCGGAGTTATGATTC 635
OY 183 AGCGGGGCGCGCGCGCGCGCTGGGCGCGCTGGGCGGCTGGTGTCTGAACCC 242
DB 636 AGCGGGGCGCGCGCGCGCGCTGGGCGCGCTGGGCGGCTGGTGTCTGAACCC 695
OY 243 GCGGGGCGCGCGCGCGCGCTGGGCGCGCTGGGCGGCTGGTGTCTGAACCC 302
DB 696 GCGTGGCGCGCGCGCGCGCTGGGCGCGCTGGGCGGCTGGTGTCTGAACCC 755
OY 303 TGAGGCTGAATCTCTCTACGCTGATCTCACTAGAGGGGAGCAAGCGCGGAGCT 362
DB 756 TGAGGCTGAATCTCTCTACGCTGATCTCACTAGAGGGGAGCAAGCGCGGAGCT 815
OY 363 GGTGGGCTGGAGAGCTGGGCGCGCTGGGCGGCTGGTGTCTGAAGAGGCT 422
DB 816 GGTGGGCTGGAGAGCTGGGCGCGCTGGGCGGCTGGTGTCTGAAGAGGCT 875
OY 423 GATGACGAGGCTGGAGAGCGCTCATGAGAGCGGCTCATGAGAGCGGCTCATGAG 482
DB 876 GATGACGAGGCTGGAGAGCGCTCATGAGAGCGGCTCATGAGAGCGGCTCATGAG 935
OY 483 GCGCGCTGTAGCTCCAGAGAGCTGGCAACCGCTGGCAGGCTTCTTGAACATTA 542
DB 936 GCGCGCTGTAGCTCCAGAGAGCTGGCAACCGCTGGCAGGCTTCTTGAACATTA 995
OY 543 TGTGCTATAGAGAGCTACCAATGAAGACCTCTGACCAATGACCTATTCCTG 602
DB 996 TGTGCTATAGAGAGCTACCAATGAAGACCTCTGACCAATGACCTATTCCTG 1055
OY 603 CCGCGCGGCTGTACCAATGAGAGCTGTCTGTGACAGCGGCTGGGCGCTGCGCT 662
DB 1056 CCGCGCGGCTGTACCAATGAGAGCTGTCTGTGACAGCGGCTGGGCGCTGCGCT 1115

OY 663 GTTCTGTGCTACAGCGCTGGGCTGCTTCTATGCTATGAGCTAGAGATGATA 722
DB 1116 CTTCTCTGTGCTACAGCGCTGGGCTGCTTCTATGCTATGAGCTAGAGATGATA 1175
OY 723 GTATCGGCTGTGGGGAGATGCGCTTCTACTCTGGGCACTTCTGCTGGACCT 782
DB 1176 GTATCGGCTGTGGGGAGATGCGCTTCTACTCTGGGCACTTCTGCTGGACCT 1235
OY 783 GCGCACTACCGCGCGCGCTGCTTCTCTTCTAGAGAGTGGTTCAGACACC 842
DB 1236 GCGCACTACCGCGCGCGCTGCTTCTCTTCTAGAGAGTGGTTCAGACACC 1295
OY 843 TGTCTCCCGCTGGTGGTCCAGAGAGCGCGGATGATGATGATGATGATGATGAT 902
DB 1296 TGTCTCCCGCTGGTGGTCCAGAGAGCGCGGATGATGATGATGATGATGATGAT 1355
OY 903 GCAGGTGCTTCTACTGCGAGGTGTCGCCAGAGAGACTTGTGCTAGTCTGGCACT 962
DB 1356 GCAGGTGCTTCTACTGCGAGGTGTCGCCAGAGAGACTTGTGCTAGTCTGGCACT 1415
OY 963 GCTGCACTGCGACCTGGCGCATGATGATGATGATGATGATGATGATGATGAT 1022
DB 1416 GCTGCACTGCGACCTGGCGCATGATGATGATGATGATGATGATGATGATGAT 1475
OY 1023 GCTGCACTGCGACCTGGCGCATGATGATGATGATGATGATGATGATGATGAT 1082
DB 1476 GCTGCACTGCGACCTGGCGCATGATGATGATGATGATGATGATGATGATGAT 1535
OY 1083 CCGTGCATGAGAGAGCGCGCATGATGATGATGATGATGATGATGATGATGAT 1142
DB 1536 CCGTGCATGAGAGAGCGCGCATGATGATGATGATGATGATGATGATGATGAT 1592
OY 1143 CGTGTGCTTCCGCTTGGAGCGCAAGGATGGGAAAGTGTGATGATGATGAT 1202
DB 1593 -----CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1605
OY 1203 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1262
DB 1606 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
OY 1263 CGTGTGCTGAGAGCGCGCATGATGATGATGATGATGATGATGATGATGAT 1322
DB 1666 TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1725
OY 1323 CTTATGACCGCTGGCGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1382
DB 1726 CTTATGACCGCTGGCGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1785
OY 1383 CTAGGCTGAGAGCGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1442
DB 1786 CTAGGCTGAGAGCGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1844
OY 1443 GGTGAGAGGTGAGAGCTTGTGGGGGACAGCGGCAAGTGAAGTCT-66GTGAG 1501
DB 1845 GGTGAGAGGTGAGAGCTTGTGGGGGACAGCGGCAAGTGAAGTGAAGTCT-66GTGAG 1904
OY 1502 AGCCAGCGTGGTGGGCGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1561
DB 1905 AGCCAGCGTGGTGGGCGCGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1960
OY 1562 ACCCCAGCAACCAATCCAAATGAAGTGAATTCCTTCCAA 1599
DB 1961 ACCCCAGCAACCAATCCAAATGAAGTGAATTCCTTCCAA 1998

RESULT 12
AF200328 1173 bp mRNA linear PRI 25-AUG-2000
LOCUS Homo sapiens sphingosine kinase (SPHK) mRNA, complete cds.
DEFINITION
ACCESSION AF200328
VERSION AF200328.1 GI:9909360
KEYWORDS

SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Mammalia: Eutheria: Primates; Catarrhini; Hominoidea: Homo.
AUTHORS Pitson, S.M., D'Andrea, R.J., Vandeleur, L., Moretti, P.A., Xia, P.,
Gamble, J.R., Vadas, M.A. and Wattenberg, B.W.
TITLE Human sphingosine kinase: purification, molecular cloning and
characterization of the native and recombinant enzymes
JOURNAL Biochem. J. 350 Pt 2, 429-441 (2000)
MEDLINE 20407120
PUBMED 10947957
REFERENCE 2 (bases 1 to 1173)
AUTHORS Pitson, S.M., D'Andrea, R.J., Vandeleur, L., Moretti, P.A., Xia, P.,
Gamble, J.R., Vadas, M.A. and Wattenberg, B.W.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-1999) Human Immunology, IMVS, Frome Road,
Adelaide, SA 5000, Australia
FEATURES
source location/Qualifiers
1..1173
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="vascular endothelial"
1..1173
/gene="SPHK"
19..1173
/gene="SPHK"
/codon_start=1
/product="sphingosine kinase"
/protein_id="A01980.1"
/db_xref="GI:909361"
translation="MDPAGGPGVLPKRCVIVILNPRGKALQLFRSHVPLAE
AIISETLMTERNRHARELVSEELGRDALVMSGDLMHEVNGMPEPDETATG
KFLCSLPAGSGNALAASLNHYAGYEQVNEELTNGTLLRLSPNLSLHTASG
LRLFSVLSLAMEFIADVDLESEKRYRLGEMFTLGTFLRLALTRYRLATLVGYV
GSKTPASPVYVQGGPDALVLEPVPVSHVTVVDEDFVLLLSHLSSEFAP
MRCAGAVMHLFVYVAGSVRAMLRFLAMEGRMEYECILYVYPVAFRLPKDG
KGRFANDSLMNSFAVQGVHNFMMVSGVPEPPSWKPOMPPEEPL"

BASE COUNT 188 a 349 c 390 g 246 t

Query Match 71.2%; Score 1137.8; DB 9; Length 1173;
Best Local Similarity 98.1%; Pred. No. 8.9e-191;
Matches 1151; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 157 GAGCCGCGGCTGAGTTATGATCCAGGCGGCGGCCCGGGGGCTGCTCCCGGCGCC 216
DB 1 GAGCCGCGGCTGAGTTATGATCCAGGCGGCGGCCCGGGGGCTGCTCCCGGCGCC 60
QY 217 TGCCGGTGTGCTGCTGTAACCCGCGCGCGGCGCAAGGCGCTTGCAGCTCTTC 276
DB 61 TGCCGGTGTGCTGCTGTAACCCGCGCGCGGCGCAAGGCGCTTGCAGCTCTTC 120
QY 277 CGGAGTCAGTGCAGCCCTTTTGGCTGAGGCTGGAATCTCTTCAAGGCTGCTCACT 336
DB 121 CGGAGTCAGTGCAGCCCTTTTGGCTGAGGCTGGAATCTCTTCAAGGCTGCTCACT 180
QY 337 GAGGCGGCGAACCAGCGCGGAGCTGTCGAGTGGGCGCGCTGGAGCGCT 396
DB 181 GAGGCGGCGAACCAGCGCGGAGCTGTCGAGTGGGCGCGCTGGAGCGCT 240
QY 397 CTGTGTGATGTCTGAGACGGGCTGATGACAGAGTGGTGAACGGGCTCATGGAGCG 456
DB 241 CTGTGTGATGTCTGAGACGGGCTGATGACAGAGTGGTGAACGGGCTCATGGAGCG 300
QY 457 CCTGACTGGAGACCGCATCCAGAAAGCCCTGTGTAGCTTCCAGAGAGGCTTGACAC 516
DB 301 CCTGACTGGAGACCGCATCCAGAAAGCCCTGTGTAGCTTCCAGAGAGGCTTGACAC 360
QY 517 GCGCTGGAGCTTCTTGAACCATTTATCTGCTGATGACAGGCTCACCAGTGAAGACCTG 576
DB 361 GCGCTGGAGCTTCTTGAACCATTTATCTGCTGATGACAGGCTCACCAGTGAAGACCTG 420

QY 577 CTGACCAACTGCACCTATTTGCTGCGCGCGGCTGCTACCATGACCTGCTGCT 636
DB 421 CTGACCAACTGCACCTATTTGCTGCGCGCGGCTGCTACCATGACCTGCTGCT 480
QY 637 CTGACACAGGCTTGGGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
DB 481 CTGACACAGGCTTGGGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 697 GCTATGTGACCTAGAGATGATGATGATGATGATGATGATGATGATGATGATG 756
DB 541 GCTATGTGACCTAGAGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 757 GGCACCTTCTGCGCTGCGAGCGCTGCGACCTACCGGCGCGCTGCTGCTGCTGCT 816
DB 601 GGCACCTTCTGCGCTGCGAGCGCTTGGCAGCTTACCGCGCGCTGCTGCTGCTGCT 660
QY 817 GTAGAGAGTGGTTCAGACACCTGCTTCCCGCTGCTGCTGCTGCTGCTGCTGCT 876
DB 661 GTAGAGAGTGGTTCAGACACCTGCTTCCCGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 877 GATGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936
DB 721 GATGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 937 GAGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996
DB 781 GAGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 997 GCACCCATGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056
DB 841 GCACCCATGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 1057 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116
DB 901 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 1117 GATGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1176
DB 961 GATGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1177 AAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1236
DB 1021 AAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1237 CACCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1296
DB 1081 CACCCAACTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1297 CAGCAGATGCGACCGCCAGAGAGCCCTTATGA 1329
DB 1141 CAGCAGATGCGACCGCCAGAGAGCCCTTATGA 1173

RESULT 13
AX127642 1155 bp DNA linear PAT 16-MAY-2001
LOCUS AX127642
DEFINITION Sequence 2 from Patent WO0131029.
ACCESSION AX127642
VERSION AX127642.1 GI:14134308
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1155)
AUTHORS Allen, J., Gosink, M., Melendez, A.J. and Takacs, L.
TITLE Human sphingosine kinase gene
JOURNAL Patent: WO 0131029-A 2 03-MAY-2001;
WARNER-LAMBERT COMPANY (US)
FEATURES
source location/Qualifiers
1..1155
/organism="Homo sapiens"
/db_xref="taxon:9606"


```

Best Local Similarity  98.2%;  Pred. No. 1.2e-184;
Matches 1115;  Conservative  0;  Mismatches  20;  Indels  0;  Gaps  0

```

Y	464	GGGAGACCGCCATCCAGAAAGCCCTGTGTAGACCTCCAGAGGCTCTGCAACACCGGCTCTG	583
Y	8	GGGAGACCGCCATCCAGAAAGCCCTGTGTAGACCTCCAGAGGCTCTGCAACACCGGCTCTG	67
Y	524	CAGCTTCCTTGACCAATTATGCTGGCTATGAGACGGTCAACATGAAAGACCTCTGACCA	583
Y	68	CAGCTTCCTTGACCAATTATGCTGGCTATGAGACGGTCAACATGAAAGACCTCTGACCA	127
Y	584	ACTGCACGCTATTTGCTGTGGCCGCCGGTGTCTACCCATGAACCTGCTGTCTGACAA	643
Y	128	ACTGCACGCTATTTGCTGTGGCCGCCGGTGTCTACCCATGAACCTGCTGTCTGACAA	187
Y	644	CGGCTTCGGGGGCTGTGGCGTTCCTGTGCTGACCTGAGCGCTGGGGCTTCATTGCTGATG	703
Y	188	CGGCTTCGGGGGCTGTGGCGTTCCTGTGCTGACCTGAGCGCTGGGGCTTCATTGCTGATG	247
Y	704	TGGACCTAGAGATGATTAAGTATGGCGCTGTGGGGGAGATGCGTTCACTCTGGGCACT	763
Y	248	TGGACCTAGAGATGATTAAGTATGGCGCTGTGGGGGAGATGCGTTCACTCTGGGCACT	307
Y	764	TCTGTGCTGTGGCAAGCCCTGGGCGACCTTACGCGGGCGAGCTGACCTCCCTGTAGGAA	823
Y	308	TCTGTGCTGTGGCAAGCCCTGGGCGACCTTACGCGGGCGAGCTGACCTCCCTGTAGGAA	367
Y	824	GAGTGGGTTTCAAGACACCTGCTTCCCGGTTGTGTGTGCACAGAGGCGGGTATGATGAC	883
Y	368	GAGTGGGTTTCAAGACACCTGCTTCCCGGTTGTGTGTGCACAGAGGCGGGTATGATGAC	427
Y	884	ACCTGGGCGACTGAGAGAGACAGTGTCTTCACTGTGACAGTGTGTGCCGACAGGACT	943
Y	428	ACCTGTGGCGACTGAGAGAGACAGTGTCTTCACTGTGACAGTGTGTGCCGACAGGACT	487
Y	944	TTGTGCTATGCTCTGGCACTGTGCACTGTGCACCTGGGCGGAGATGATGTTTCTGACCCA	1003
Y	488	TTGTGCTATGCTCTGGCACTGTGCACTGTGCACCTGGGCGGAGATGATGTTTCTGACCCA	547
Y	1004	TGGGCGCGTGTGAGCTGTGGCGTCAATGATCTTCTTACGTGGCGGGGAGTGTCTGTG	1063
Y	548	TGGGCGCGTGTGAGCTGTGGCGTCAATGATCTTCTTACGTGGCGGGGAGTGTCTGTG	607
Y	1064	CCATGCTGTGGCGCTTCTTCCGAGGATGAGAGAGGCGAGCAATGAGATATGATGCC	1123
Y	608	CCATGCTGTGGCGCTTCTTCCGAGGATGAGAGAGGCGAGCAATGAGATATGATGCC	667
Y	1124	CCCTACTGTGATATGTGCCCGCGTGTGCGCTTCCGCTTGGAGCCCAAGATGGGAAAGTG	1183
Y	668	CCCTACTGTGATATGTGCCCGCGTGTGCGCTTCCGCTTGGAGCCCAAGATGGGAAAGTG	727
Y	1184	TGTTTGCAGTGAGATGGGGAATTGATGATGTAAGGAGGCGGTGAGAGGCCAGGTCACCCA	1243
Y	728	TGTTTGCAGTGAGATGGGGAATTGATGATGTAAGGAGGCGGTGAGAGGCCAGGTCACCCA	787
Y	1244	ACTACTTGTGATGATGAGCGGTGTGCGTGTGAGGCCCGCGCCACGCTGGAAAGCCCAAGCA	1303
Y	788	ACTACTTGTGATGATGAGCGGTGTGCGTGTGAGGCCCGCGCCACGCTGGAAAGCCCAAGCA	847
Y	1304	TGCCACCGCCGAAAGAGCCCTTATGACACCTGGGCGCGGCTGTGCTTATGTTTACATTG	1363
Y	848	TGCCACCGCCGAAAGAGCCCTTATGACACCTGGGCGCGGCTGTGCTTATGTTTACATTG	907
Y	1364	CAGGACCCCTTCTCCTTCCCTAGAGGCTGCAAGGGCTGTGCACAGCTCTGTGAGGGGTGA	1423
Y	908	CAGGACCCCTTCTCCTTCCCTAGAGGCTGCAAGGGCTGTGCACAGCTCTGTGAGGGGTGA	967
Y	1424	GGAGACTCTCTGAGGAAGGGTGAAGAGTGAAGAGTATGCTTTGGGGGGACAGGCGCAA	1483
Y	968	GGAGACTCTCTGAGGAAGGGTGAAGAGTGAAGAGTATGCTTTGGGGGGACAGGCGCAA	1027
Y	1484	ATGAAGTCTGTGGGTGAGAGGCCACAGTGTGCTGGGCCACAGTGTCTATGTAAGGCTTCTA	1544

Db 1028 ATGAAGTCTGGGTCAGAGCCCAAGCTGCGCTGGGCCCAAGTGGCTATGTAAGAGCCCTCTCA 1087

QY 1544 GTTTGTCTGAGAACCCCAACCCACAGAACCAATCCAAATAAAGTGACATTCCCA 1598

Db 1088 GTTTGTCTGAGAACCCCAACCCACAGAACCAATCCAAATAAAGTGACATTCCCA 1142

AC068145/c	183443 bp	DNA	linear	HTG 04-MAY-2001
LOCUS				
DEFINITION				
AC068145				
Homo sapiens chromosome 17 clone CTF-2531H				
SEQUENCE, 48 unordered clones.				
				map 17, WORKING DRAFT

ACCESSION	AC068145
VERSION	AC068145.3
KEYWORDS	GI:9966946
	HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE	ORGANISM
Homo sapiens.	
Homo sapiens	
Standard. Creatista: Vertebrata: Euteleostomi:	

REFERENCE
1 (bases 1 to 183443)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Eukaryota; Metazoa; Chordata; Cnidaria; Vertebrata; Mammalia; Primates; Catarrhini; Hominoidea; Homo.
Nuchbaum C and Landfer E.
1 (bases 1 to 183443)

AUTHORS	TITLE	JOURNAL	REFERENCE
Britten, B., Linton, L., Nusbaum, C., Lander, E., Abramson, H., Allen, N., Batten, V., Beda, E., Davis, N.	Homo sapiens chromosome 17, clone CTD-2531H7	Unpublished	2 (bases 1 to 183443)

Anderson, S., Balow III, C., Balaia, N., Dabrowski, J.,
Boguslawski, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campanelli, A., Castle, A., Choepey, Y., Colangelo, M., Collins, S.,

Campopiano, T., Cooke, P., Deatellano, K., Dewar, R., Diaz, J.S.,
Collamore, A., Domino, M., Doyle, M., Ferrelita, P., Fitzhugh, W., Gage, D.,
Dodge, S., Domino, M., Doyle, M., Ferrelita, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glende, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Kamran, I., Karpas, A., Lee, C., Looze, C.

Howland, J. C., Iliev, I., Johnson, K., Jones, C., Karmali, M. A., Khatami, N., Klein, J., Lacroix, R., Landers, T., Lehoczký, I., Levine, R., Lieu, C., Liu, G., Locke, K., Lamacz, R., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheeters, R., Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.

Murphy, T., Maylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pislani, C., Pollara, V., Raymond, C., Riley, R., Rogov, R., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stachowicz, N., Stachowicz, N., Subramanian, A., Talamas, J.,

STREIBER
Direct Submission
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Stange-Thomann, N., Stojanovic, N., Subramanian, S.,
Suzuki, T.

JOURNAL Submitted (29-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 5, 2000 this sequence version replaced gi:7960325.
COMMENT All repeats were identified using RepeatMasker:

Email: A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB

Center code: W1100
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L10026

```

Center clone name: 2531_H_7
----- Summary Statistics -----
Sequencing vector: M13; M77815; 99% of reads
Sequencing vector: Plasmid; n/a; <0.1% of reads
Sequencing vector: chemistrv: n/a; terminator Big Dye; 100% of reads

```

```

reads
0.756/253414204VChrom122.1.f - 230 consensus
Assembly program: Phrap: version 0.960731
Consensus quality: 152281 bases at least Q40
Consensus quality: 167680 bases at least Q30

```

Consensus quality: 174098 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 178743; sum-of-contents
Quality coverage: 2.9 in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1059: contlg of 1059 bp	in length
1060	1159: gap of 100 bp	
1160	2777: contlg of 1618 bp	in length
2778	2877: gap of 100 bp	
2878	4283: contlg of 1406 bp	in length
4284	4383: gap of 100 bp	
4384	5447: contlg of 1064 bp	in length
5448	5547: gap of 100 bp	
5548	6766: contlg of 1228 bp	in length
6777	6876: gap of 100 bp	
6877	8215: contlg of 1339 bp	in length
8216	8315: gap of 100 bp	
8316	9961: contlg of 1646 bp	in length
9962	10061: gap of 100 bp	
10062	12224: contlg of 2163 bp	in length
12225	12324: gap of 100 bp	
12325	13794: contlg of 1470 bp	in length
13795	13894: gap of 100 bp	
13895	16857: contlg of 2963 bp	in length
16858	16957: gap of 100 bp	
16958	18266: contlg of 1309 bp	in length
18267	18366: gap of 100 bp	
18367	20190: contlg of 1824 bp	in length
20191	20290: gap of 100 bp	
20291	21460: contlg of 1170 bp	in length
21461	21560: gap of 100 bp	
21561	23407: contlg of 1847 bp	in length
23408	23507: gap of 100 bp	
23508	23455: contlg of 1948 bp	in length
25456	25555: gap of 100 bp	
25556	27331: contlg of 1776 bp	in length
27332	27431: gap of 100 bp	
27432	30406: contlg of 2975 bp	in length
30407	30506: gap of 100 bp	
30507	32805: contlg of 2299 bp	in length
32806	32905: gap of 100 bp	
32906	33659: contlg of 2754 bp	in length
33660	35759: gap of 100 bp	
35760	37974: contlg of 2215 bp	in length
37975	38074: gap of 100 bp	
38075	40521: contlg of 2447 bp	in length
40522	40621: gap of 100 bp	
40622	43219: contlg of 2598 bp	in length
43220	43319: gap of 100 bp	
43320	46161: contlg of 2842 bp	in length
46162	46261: gap of 100 bp	
46262	48463: contlg of 2202 bp	in length
48464	48563: gap of 100 bp	
48564	51275: contlg of 2712 bp	in length
51276	51375: gap of 100 bp	
51376	53522: contlg of 2147 bp	in length
53523	53622: gap of 100 bp	
53623	55955: contlg of 3333 bp	in length
55956	57055: gap of 100 bp	
57056	59555: contlg of 2500 bp	in length
59556	59655: gap of 100 bp	
59656	62713: contlg of 3058 bp	in length
62714	62813: gap of 100 bp	
62814	66519: contlg of 3706 bp	in length
66520	66619: gap of 100 bp	
66620	66930: contlg of 2311 bp	in length
66931	69030: gap of 100 bp	
69031	72724: contlg of 3694 bp	in length
72725	72824: gap of 100 bp	
72825	76092: contlg of 3268 bp	in length

FEATURES	SOURCE
* 76093 76192: gap of	100 bp
* 76193 80143: contig of 3951 bp	in length
* 80144 80243: gap of	100 bp
* 80244 84957: contig of 4714 bp	in length
* 84958 85057: gap of	100 bp
* 85058 88424: contig of 3367 bp	in length
* 88425 88524: gap of	100 bp
* 88525 93428: contig of 4904 bp	in length
* 93429 93528: gap of	100 bp
* 93529 97177: contig of 3649 bp	in length
* 97178 97277: gap of	100 bp
* 97278 102713: contig of 3436 bp	in length
* 102714 102813: gap of	100 bp
* 102814 109141: contig of 6328 bp	in length
* 109142 109241: gap of	100 bp
* 109242 114513: contig of 5372 bp	in length
* 114514 114613: gap of	100 bp
* 114614 124525: contig of 9912 bp	in length
* 124526 124625: gap of	100 bp
* 124626 132364: contig of 7739 bp	in length
* 132365 132464: gap of	100 bp
* 132465 139488: contig of 7024 bp	in length
* 139489 139588: gap of	100 bp
* 139589 148734: contig of 9146 bp	in length
* 148735 148834: gap of	100 bp
* 148835 159465: contig of 10631 bp	in length
* 159466 159565: gap of	100 bp
* 159566 172685: contig of 13120 bp	in length
* 172686 172785: gap of	100 bp
* 172786 183443: contig of 10658 bp	in length
Location/Qualifiers	
1..183443	

Query Match	Similarity	63.6%	Score	1017.4	DB 2:	Length	183443
Best Local	Similarity	98.0%	Pred	No. 6	9e-170:		
Matches	1030:	Conservative	0:	Mismatches	21:	Indels	0:
						Gaps	0:
QY	548	GCTATGAGCAGGCTCACCAATGTAAGACCTCTCTGACCACTGCACGCTATTGCTGTGCGGCC	607				
Db	134217	GCTATGAGCAGGCTCACCAATGTAAGACCTCTCTGACCACTGCACGCTATTGCTGTGCGGCC	134158				
QY	608	CGGTCCTGTACACCCATGGAACCTGCTGTCTCTGCACACGGCTTCGGGGCTCGCGCTCGTTCT	667				

```
Db 134157 GGGTCTGTACACCATGATACCTGCTCTGTGCACAGCGCTCGGGGCTGCGCTCTCTT 134098
QY 668 CTGTGCTCAGCCTTGCGCTTGCGGCTTCATTCGTATGTGACCTAGAGAGTATAGTATC 727
Db 134097 CTGTGCTCAGCCTTGCGCTTGCGGCTTCATTCGTATGTGACCTAGAGAGTATAGTATC 134038
QY 728 GGGCTCTGGGGAGATCGCTTCACCTGTGCGGCACTTCCTGCTGCGAGCCCTGCGCA 787
Db 134037 GGGCTCTGGGGAGATCGCTTCACCTGTGCGGCACTTCCTGCTGCGAGCCCTGCGCA 133978
QY 788 CTTACCGCGGCGGCTGCTGCTACCTCCCTGTAGAGAGTGGGTTTCAAGACACTGCTT 847
Db 133977 CTTACCGCGGCGGCTGCTGCTACCTCCCTGTAGAGAGTGGGTTTCAAGACACTGCTT 133918
QY 848 CCCCCGTTTGTGTCCAGAGAGGCCGGTATGATGCACACTGCTGCCACTGAGAGAGCAGG 907
Db 133917 CCCCCGTTTGTGTCCAGAGAGGCCGGTATGATGCACACTGCTGCCACTGAGAGAGCAGG 133858
QY 908 TGCCTTCTACTGTGCGAGGTGGTCCCGACGAGACTTGTGCTAGTCTCTGGCACTGCTGC 967
Db 133857 TGCCCTCTACTGTGCGAGGTGGTCCCGACGAGACTTGTGCTAGTCTCTGGCACTGCTGC 133798
QY 968 ACTGGCACTTGGCGAGTGAATGTTTGTGTCACCATGGGCGCTGTGACAGTGGCTCA 1027
Db 133797 ACTGGCACTTGGCGAGTGAATGTTTGTGTCACCATGGGCGCTGTGACAGTGGCTCA 133738
QY 1028 TGCATCTGTTCTAGCTGCGGGCGGGAGTGTCTGCTGCCATGCTGCTGCGCTTCTCTGG 1087
Db 133737 TGCATCTGTTCTAGCTGCGGGCGGGAGTGTCTGCTGCCATGCTGCTGCGCTTCTCTGG 133678
QY 1088 CCATGGAGAGGGCGAGCATATGAGTATGATGCCCTACTTGGTATATGTGCCCGTGG 1147
Db 133677 CCATGGAGAGGGCGAGCATATGAGTATGATGCCCTACTTGGTATATGTGCCCGTGG 133618
QY 1148 TCGCCTTCGCTTGGAGCCCAAGATGGGAAAGTGTGTTGACAGTGGATGGGAATTGA 1207
Db 133617 TCGCCTTCGCTTGGAGCCCAAGATGGGAAAGTGTGTTGACAGTGGATGGGAATTGA 133558
QY 1208 TGGTTAGCGAGGCGCTGCAAGGCGCAAGTGTGCAACCACTACTTCTGATGTGCAAGCGTT 1267
Db 133557 TGGTTAGCGAGGCGCTGCAAGGCGCAAGTGTGCAACCACTACTTCTGATGTGCAAGCGTT 133498
QY 1268 GCGTGGAGGCGCCGCGCCAGCTGGAAGCCCAAGATGCCACCGCCAGAGAGGCCCTTAT 1327
Db 133497 GCGTGGAGGCGCCGCGCCAGCTGGAAGCCCAAGATGCCACCGCCAGAGAGGCCCTTAT 133438
QY 1328 GACCCCTGGGCGCGCGCTGTGCTTAGTGTCTACTTGCAGAGACCTTCTCTTCCCTTAGG 1387
Db 133437 GACCCCTGGGCGCGCGCTGTGCTTAGTGTCTACTTGCAGAGACCTTCTCTTCCCTTAGG 133378
QY 1388 GCTGGCAGGCGCTGTCCACAGCTCCTGTGGGGGTGGAGAGACTCCTGAGAGAGGTGA 1447
Db 133377 GCTGGCAGGCGCTGTCCACAGCTCCTGTGGGGGTGGAGAGACTCCTGAGAGAGGTGA 133318
QY 1448 GAAGTGAAGGCTATGCTTGGGGGACAGGCCAAGATGAAGTCTGGGTGAGAGCCCA 1507
Db 133317 GAAGTGAAGGCTATGCTTGGGGGACAGGCCAAGATGAAGTCTGGGTGAGAGCCCA 133258
QY 1508 GCTGGCTGGGCGCAGTGCCTATGTAGAGCCCTTCTAGTGTGTTCTGAGAGCCCGCACCCCA 1567
Db 133257 GCTGGCTGGGCGCAGTGCCTATGTAGAGCCCTTCTAGTGTGTTCTGAGAGCCCGCACCCCA 133198
QY 1568 CGAACCATAATCCAAATTAAGTACATTCCCA 1598
Db 133197 CGAACCATAATCCAAATTAAGTACATTCCCA 133167
```